Mouse Dnmt3a DNA sequence

1 GAATTCCGGC CTGCTGCCGG GCCGCCCGAC CCGCCGGGCC ACACGGCAGA

51 GCCGCCTGAA GCCCAGCGCT GAGGCTGCAC TTTTCCGAGG GCTTGACATC

101 AGGGTCTATG TTTAAGTCTT AGCTCTTGCT TACAAAGACC ACGGCAATTC

151 CTTCTCTGAA GCCCTCGCAG CCCCACAGCG CCCTCGCAGC CCCAGCCTGC

201 CGCCTACTGC CCAGCAATGC CCTCCAGCGG CCCCGGGGAC ACCAGCAGCT

251 CCTCTCTGGA GCGGGAGGAT GATCGAAAGG AAGGAGAGGA ACAGGAGGAG

301 AACCGTGGCA AGGAAGAGCG CCAGGAGCCC AGCGCCACGG CCCGGAAGGT

351 GGGGACGCCT GGCCGGAAGC GCAAGCACCC ACCGGTGGAA AGCAGTGACA

401 CCCCCAAGGA CCCAGCAGTG ACCACCAAGT CTCAGCCCAT GGCCCAGGAC

451 TCTGGCCCCT CAGATCTGCT ACCCAATGGA GACTTGGAGA AGCGGAGTGA

501 ACCCCAACCT GAGGAAGGGA GCCCAGCTGC AGGGCAGAAG GGTGGGGCCC

551 CAGCTGAAGG AGAGGGAACT GAGACCCCAC CAGAAGCCTC CAGAGCTGTG

601 GAGAATGGCT GCTGTGTGAC CAAGGAAGGC CGTGGAGCCT CTGCAGGAGA

651 GGGCAAAGAA CAGAAGCAGA CCAACATCGA ATCCATGAAA ATGGAGGGCT

701 CCCGGGGCCG ACTGCGAGGT GGCTTGGGCT GGGAGTCCAG CCTCCGTCAG

751 CGACCCATGC CAAGACTCAC CTTCCAGGCA GGGGACCCCT ACTACATCAG

801 CAAACGGAAA CGGGATGAGT GGCTGGCACG TTGGAAAAGG GATGCTGAGA

851 AGAAAGCCAA GGTAATTGCA GTAATGAATG CTGTGGAAGA GAACCAGGCC

901 TCTGGAGAGT CTCAGAAGGT GGAGGAGGCC AGCCCTCCTG CTGTGCAGCA

951 GCCCACGGAC CCTGCTTCTC CGACTGTGGC CACCACCCCT GAGCCAGTAG

1001 GAGGGGATGC TGGGGACAAG AATGCTACCA AAGCACCCGA CGATGAGCCT

1051 GAGTATGAGG ATGGCCCGGG CTTTGGCATT GGAGAGCTGG TGTGGGGGAA

1101 ACTTCGGGGT TTCTCTTGGT GGCCAGGCCG AATTGTGTCT TGGTGGATGA

FIG. 1A-1

1151 CAGGCCGGAG CCGAGCAGCT GAAGGCACTC GCTGGGTCAT GTGGTTCGGA

1201 GATGGCAAGT TCTCAGTGGT GTGTGTGGAG AAGCTCATGC CGCTGAGCTC

1251 CTTCTGCAGT GCATTCCACC AGGCCACCTA CAACAAGCAG CCCATGTACC

1301 GCAAAGCCAT CTACGAAGTC CTCCAGGTGG CCAGCAGCCG TGCCGGGAAG

1351 CTGTTTCCAG CTTGCCATGA CAGTGATGAA AGTGACAGTG GCAAGGCTGT

1401 GGAAGTGCAG AACAAGCAGA TGATTGAATG GGCCCTCGGT GGCTTCCAGC

1451 CCTCGGGTCC TAAGGGCCTG GAGCCACCAG AAGAAGAGAA GAATCCTTAC

1501 AAGGAAGTTT ACACCGACAT GTGGGTGGAG CCTGAAGCAG CTGCTTACGC

1551 CCCACCCCCA CCAGCCAAGA AACCCAGAAA GAGCACAACA GAGAAACCTA

1601 AGGTCAAGGA GATCATTGAT GAGCGCACAA GGGAGCGGCT GGTGTATGAG

1651 GTGCGCCAGA AGTGCAGAAA CATCGAGGAC ATTTGTATCT CATGTGGGAG

1701 CCTCAATGTC ACCCTGGAGC ACCCATTCTT CATTGGAGGC ATGTGCCAGA

1751 ACTGTAAGAA CTGCTTCTTG GAGTGTGCTT ACCAGTATGA CGACGATGGG

1801 TACCAGTCCT ATTGCACCAT CTGCTGTGGG GGGCGTGAAG TGCTCATGTG

1851 TGGGAACAAC AACTGCTGCA GGTGCTTTTG TGTCGAGTGT GTGGATCTCT

1901 TGGTGGGGCC AGGAGCTGCT CAGCCAGCCA TTAAGGAAGA CCCCTGGAAC

1951 TGCTACATGT GCGGGCATAA GGGCACCTAT GGGCTGCTGC GAAGACGGGA

2001 AGACTGGCCT TCTCGACTCC AGATGTTCTT TGCCAATAAC CATGACCAGG

2051 AATTTGACCC CCCAAAGGTT TACCCACCTG TGCCAGCTGA GAAGAGGAAG

2101 CCCATCCGCG TGCTGTCTCT CTTTGATGGG ATTGCTACAG GGCTCCTGGT

2151 GCTGAAGGAC CTGGGCATCC AAGTGGACCG CTACATTGCC TCCGAGGTGT

2201 GTGAGGACTC CATCACGGTG GGCATGGTGC GCCACCAGGG AAAGATCATG

2251 TACGTCGGGG ACGTCCGCAG CGTCACACAG AAGCATATCC AGGAGTGGGG

2301 CCCATTcGAC cTGGTGATTG GAGGCAGTCC CTGCAATGAC CTcTCCATTG

FIG. 1A-2

ICET20 08 02 00

2351 TCAACCCTGC CCGCAAGGGA CTTTATGAGG GTACTGGCCG CCTCTTCTTT

2401 GAGTTCTACC GCCTCCTGcA TGATGCGCGG CCCAAGGAGG GAGATGATCG

2451 CCCCTTCTTC TGGCTCTTTG AGAATGTGGT GGCCATGGGC GTTAGTGACA

2501 AGAGGGACAT CTCGCGATTT CTTGAGTCTA ACCCCGTGAT GATTGACGCC

2551 AAAGAAGTGT CTGCTGCACA CAGGGCCCGT TACTTCTGGG GTAACCTTCC

2601 TGGCATGAAC AGGCCTTTGG CATCCACTGT GAATGATAAG CTGGAGCTGC

2651 AAGAGTGTCT GGAGCACGGC AGAATAGCCA AGTTCAGCAA AGTGAGGACC

2701 ATTACCACCA GGTCAAACTC TATAAAGCAG GGCAAAGACC AGCATTTCCC

2751 CGTCTTCATG AACGAGAAGG AGGACATCCT GTGGTGCACT GAAATGGAAA

2801 GGGTGTTTGG CTTCCCCGTC CACTACACAG ACGTCTCCAA CATGAGCCGC

2851 TTGGCGAGGC AGAGACTGCT GGGCCGATCG TGGAGCGTGC CGGTCATCCG

2901 CCACCTCTTC GCTCCGCTGA AGGAATATTT TGCTTGTGTG TAAGGGACAT

2951 GGGGGCAAAC TGAAGTAGTG ATGATAAAAA AGTTAAACAA ACAAACAAAC

3001 AAAAAACAAA ACAAAACAAT AAAACACCAA GAACGAGAGG ACGGAGAAAA

3051 GTTCACCACC CAGAAGAGAA AAAGGAATTT AAAGCAAACC ACAGAGGAGG

3101 AAAACGCCGG AGGGCTTGGC CTTGCAAAAG GGTTGGACAT CATCTCCTGA

3151 GTTTTCAATG TTAACCTTCA GTCCTATCTA AAAAGCAAAA TAGGCCCCTC

3201 CCCTTCTTCC CCTCCGGTCC TAGGAGGCGA ACTTTTTGTT TTCTACTCTT

3251 TTTCAGAGGG GTTTTCTGTT TGTTTGGGTT TTTGTTTCTT GCTGTGACTG

3301 AAACAAGAGA GTTATTGCAG CAAAATCAGT AACAACAAAA AGTAGAAATG

3351 CCTTGGAGAG GAAAGGGAGA GAGGGAAAAT TCTATAAAAA CTTAAAATAT

3401 TGGTTTTTTT TTTTTTTCCT TTTCTATATA TCTCTTTGGT TGTCTCTAGC

3451 CTGATCAGAT AGGAGCACAA ACAGGAAGAG AATAGAGACC CTCGGAGGCA

3501 GAGTCTCCTC TCCCACCCCC CGAGCAGTCT CAACAGCACC ATTCCTGGTC

FIG. 1A—3

T05T20.98DD2760

```
3551  ATGCAAAACA GAACCCAACT AGCAGCAGGG CGCTGAGAGA ACACCACACC
3601  AGACACTTTC TACAGTATTT CAGGTGCCTA CCACACAGGA AACCTTGAAG
3651  AAAACCAGTT TCTAGAAGCC GCTGTTACCT CTTGTTTACA GTTTATATAT
3701  ATATGATAGA TATGAGATAT ATATATATAA AAGGTACTGT TAACTACTGT
3751  ACATCCCGAC TTCATAATGG TGCTTTCAAA ACAGCGAGAT GAGCAAAGAC
3801  ATCAGCTTCC GCCTGGCCCT CTGTGCAAAG GGTTTCAGCC CAGGATGGGG
3851  AGAGGGGAGC AGCTGGAGGG GGTTTTAACA AACTGAAGGA TGACCCATAT
3901  CACCCCCCAC CCCTGCCCCA TGCCTAGCTT CACCTGCCAA AAAGGGGCTC
3951  AGCTGAGGTG GTCGGACCCT GGGGAAGCTG AGTGTGGAAT TTATCCAGAC
4001  TCGCGTGCAA TAACCTTAGA ATATGAATCT AAAATGACTG CCTCAGAAAA
4051  ATGGCTTGAG AAAACATTGT CCCTGATTTT GAATTCGTCA GCCACGTTGA
4101  AGGCCCCTTG TGGGATCAGA AATATTCCAG AGTGAGGGAA AGTGACCCGC
4151  CATTAACCCC NCCTGGAGCA AATAAAAAAA CATACAAAAT GT
```

FIG. 1A–4

AMENDED SHEET

Mouse Dnmt3b1 DNA Sequence

1 GAATTCCGGG CGCCGGGGTT AAGCGGCCCA AGTAAACGTA GCGCAGCGAT

51 CGGCGCCGGA GATTCGCGAA CCCGACACTC CGCGCCGCCC GCCGGCCAGG

101 ACCCGCGGCG CGATCGCGGC GCCGCGCTAC AGCCAGCCTC ACGACAGGCC

151 CGCTGAGGCT TGTGCCAGAC CTTGGAAACC TCAGGTATAT ACCTTTCCAG

201 ACGCGGGATC TCCCCTCCCC CATCCATAGT GCCTTGGGAC CAAATCCAGG

251 GCCTTCTTTC AGGAAACAAT GAAGGGAGAC AGCAGACATC TGAATGAAGA

301 AGAGGGTGCC AGCGGGTATG AGGAGTGCAT TATCGTTAAT GGGAACTTCA

351 GTGACCAGTC CTCAGACACG AAGGATGCTC CCTCACCCCC AGTCTTGGAG

401 GCAATCTGCA CAGAGCCAGT CTGCACACCA GAGACCAGAG GCCGCAGGTC

451 AAGCTCCCGG CTGTCTAAGA gGGAGGTCTC CAgCCTTCTG AATTACACGC

501 AGGACATGAC AGGAGATGGA GACAGAGATG ATGAAGTAGA TGATGGGAAT

551 GGCTCTGATA TTCTAATGCC AAAGCTCACC CGTGAGACCA AGGACACCAG

601 GACGCGCTCT GAAAGCCCGG CTGTCCGAAC CCGACATAgC AATGGGACCT

651 CCAGCTTGGA CAGGCAAAGA GCCTCCCCCA gAATCACCCG AGGTCGGCAG

701 GGCCGCCACC ATGTGCAGGA GTACCCTGTG GAGTTTCCGG CTACCAGGTC

751 TCGGAGACGT CGAGCATCGT CTTCAGCAAG CACGCCATGG TCATCCCCTG

801 CCAGCGTCGA CTTCATGGAA GAAGTGACAC CTAAGAGCGT CAGTACCCCA

851 TCAGTTGACT TGAGCCAGGA TGGAGATCAG GAGGGTATGG ATACCACACA

901 GGTGGATGCA GAGAGCATAT ATGGAgACAG CACAGAGTAT CAgGATGATA

951 AAGAGTTTGG AATAGGTGAC CTCGTGTGGG GAAAGATCAA GGGCTTCTCC

1001 TGGTGGCCTG CCATGGTGGT GTCCTGGAAA GCCACCTCCA AgCGACAGGC

FIG. 1B-1

AMENDED SHEET

TDE:\50-98002\50

1051 CATGCCCGGA ATGCGCTGGG TACAGTGGTT TGGTGATGGC AAGTTTTCTG

1101 AGATCTCTGC TGACAAACTG GTGGCTCTGG GGCTGTTCAG CCAGCACTTT

1151 AATCTGGCTA CCTTCAATAA GCTGGTTTCT TATAGGAAGG CCATGTACCA

1201 CACTCTGGAG AAAGCCAGGG TTCGAGCTGG CAAGACCTTC TCCAGCAGTC

1251 CTGGAGAGTC ACTGGAGGAC CAGCTGAAGC CCATGCTGGA GTGGGCCCAC

1301 GCTGGCTTCA AGCCTACTGG GATCGAGGGC CTCAAACCCA ACAAGAAGCA

1351 ACCAGTGGTT AATAAGTCGA AGGTGCGTCG TTCAGACAGT AGGAACTTAG

1401 AACCCAGGAG ACGCGAGAAC AAAAGTCGAA GACGCACAAC CAATGACTCT

1451 GCTGCTTCTG AGTCCCCCCC ACCCAAGCGC CTCAAGACAA ATAGCTATGC

1501 CGGGAAGGAC CGAGGGGAGG ATGAGGAGAG CCGAGAACGG ATGGCTTCTG

1551 AAGTCACCAA CAACAAGGGC AATCTGGAAG ACCGCTGTTT GTCCTGTGGA

1601 AAGAAGAACC CTGTGTCCTT CCACCCCCTC TTTGAGGGTG GGCTCTGTCA

1651 GAGTTGCCGG GATCGCTTCC TAGAGCTCTT CTACATGTAT GATGAGGACG

1701 GCTATCAGTC CTACTGCACC GTGTGCTGTG AGGGCCGTGA ACTGCTGCTG

1751 TGCAGTAACA CAAGCTGCTG CAGATGCTTC TGTGTGGAGT GTCTGGAGGT

1801 GCTGGTGGGC GCAGGCACAG CTGAGGATGC CAAGCTGCAG GAACCCTGGA

1851 GCTGCTATAT GTGCCTCCCT CAGCGCTGCC ATGGGGTCCT CCGACGCAGG

1901 AAAGATTGGA ACATGCGCCT GCAAGACTTC TTCACTACTG ATCCTGACCT

1951 GGAAGAATTT GAGCCACCCA AGTTGTACCC AGCAATTCCT GCAGCCAAAA

2001 GGAGGCCCAT TAGAGTCCTG TCTCTGTTTG ATGGAATTGC AACGGGGTAC

2051 TTGGTGCTCA AGGAGTTGGG TATTAAAGtG GAAAAGTACA TTGCCTCCGA

2101 AGTCTGTGCA GAGTCCATCG CTGTGGGAAC TGTTAAGCAT GAAGGCCAGA

2151 TCAAATATGT CAATGACGTC CGGAAAATCA CCAAGAAAAA TATTGAAGAG

2201 TGGGGCCCGT TCGACTTGGT GATTGGTGGA AGCCCATGCA ATGATCTCTC

FIG. 1B–2

FDE 20 '98 02:50

2251 TAACGTCAAT CCTGCCCGCA AAGGTTTATA TGAGGGCACA GGAAGGCTCT

2301 TCTTCGAGTT TTACCACTTG CTGAATTATA CCCGCCCCAA GGAGGGCGAC

2351 AACCGTCCAT TCTTCTGGAT GTTCGAGAAT GTTGTGGCCA TGAAAGTGAA

2401 TGACAAGAAA GACATCTCAA GATTCCTGGC ATGTAACCCA GTGATGATCC

2451 ATGCCATCAA GGTGTCTGCT GCTCACAGGG CCCGGTACTT CTGGGGTAAC

2501 CTACCCGGAA TGAACAGGCC CGTGATGGCT TCAAAGAATG ATAAGCTCGA

2551 GCTGCAGGAC TGCCTGGAGT TCAGTAGGAC AGCAAAGTTA AAGAAAGTGC

2601 AGACAATAAC CACCAAGTCG AACTCCATCA GACAGGGCAA AAACCAGCTT

2651 TTCCCTGTAG TCATGAATGG CAAGGACGAC GTTTTGTGGT GCACTGAGCT

2701 CGAAAGGATC TTCGGCTTCC CTGCTCACTA CACGGACGTG TCCAACATGG

2751 GCCGCGGCGC CCGTCAGAAG CTGCTGGGCA GGTCCTGGAG TGTACCGGTC

2801 ATCAGACACC TGTTTGCCCC CTTGAAGGAC TACTTTGCCT GTGAATAGTT

2851 CTACCCAGGA CTGCGGAGCT CTCGGTCAGA GCCAGTGCCC AGAGTCACCC

2901 CTCCCTGAAG GCACCTCACC TGTCCCCTTT TTAGCTCACC TGTGTGGGGC

2951 CTCACATCAC TGTACCTCAG CTTTCTCCTG CTCAGTGGGA GCAGAGCCTC

3001 CTGGCCCTTG CAGGGGAGCC CCGGTGCTCC CTCCGTGTGC ACAGCTCAGA

3051 CCTGGCTGCT TAGAGTAGCC CGGCATGGTG CTCATGTTCT CTTACCCTGA

3101 AACTTTAAAA CTTGAAGTAG GTAGTAAGAT GGCTTTCTTT TACCCTCCTG

3151 AGTTTATCAC TCAGAAGTGA TGGCTAAGAT ACCAAAAAAA CAAACAAAAA

3201 CAGAAACAAA AAACAAAAAA AAACCTCAAC AGCTCTcTTA GTACTCAGGT

3251 TCATGCTGCA AAATCACTTG AGATTTTGTT TTTAAGTAAC CCGTGcTCcA

3301 CATTTGCTGG AGGATGCTAT TGTGAATGTG GGCTCAGATG AGCAAGGTCA

3351 AGGGGCCAAA AAAAATTCCC CCTCTCCCCC CAGGAGTATT TGAAGATGAT

3401 GTTTATGGTT TAAGTCTTCC TGGCACCTTC CCCTTGCTTT GGTACAAGGG

FIG.1B-3

FOETJ0 '96002460

```
3451  CTGAAGTCCT GTTGGTCTTG TAGCATTTCC CAGGATGATG ATGTCAGCAG
3501  GGATGACATC ACCACCTTTA GGGCTTTTCC CTGGCAGGGG CCCATGTGGC
3551  TAGTCCTCAC GAAGACTGGA GTAGAATGTT TGGAGCTCAG GAAGGGTGGG
3601  TGGAGTGGCC CTCTTCCAGG TGTGAGGGAT ACGAAGGAGG AAGCTTAGGG
3651  AAATCCATTC CCCACTCCCT CTTGCCAAAT GAGGGCCCA GTCCCCAACA
3701  GCTCAGGTCC CCAGAACCCC CTAGTTCCTC ATGAGAAGCT AGGACCAGAA
3751  GCACATCGTT CCCCTTATCT GAGCAGTGTT TGGGGAACTA CAGTGAAAAC
3801  CTTCTGGAGA TGTTAAAAGC TTTTTACCCC ACGATAGATT GTGTTTTTAA
3851  GGGGTGCTTT TTTTAGGGGC ATCACTGGAG ATAAGAAAGC TGCATTTCAG
3901  AAATGCCATC GTAATGGTTT TTAAACACCT TTTACCTAAT TACAGGTGCT
3951  ATTTTATAGA AGCAGACAAC ACTTCTTTTT ATGACTCTCA GACTTCTATT
4001  TTCATGTTAC CATTTTTTTT GTAACTCGCA AGGTGTGGGC TTTTGTAACT
4051  TCACAGGTGT GGGGAGAGAC TGCCTTGTTT CAACAGTTTG TCTCCACTGG
4101  TTTCTAATTT TTAGGTGCAA AGATGACAGA TGCCCAGAGT TTACCTTTCT
4151  GGTTGATTAA AGTTGTATTT CTCTAAAAAA AAAAAAAAAA AAAAA
```

FIG. 1B–4

FORM PCT/RO/134 (JULY 1998)

09/723086

Human DNMT3A DNA Sequence


```
1     GGCCGGCGTC GACCGACAGC GAGCGGAGGG AGGGAGCGAG CGAGCGACCA
51    GCAGCGGCCG GGAGGGAGGG AGGGCGCGCG GGCGGCGGCG GCGGCGAGAG
101   CAGAGGACGA GCCGGGACGC GGCGCCGCGG CACCAGGGCG CGCAGCCGGG
151   CCGGCCCGAC CCCACCGGCC ATACGGTGGA GCCATCGAAG CCCCCACCCA
201   CAGGCTGACA GAGGCACCGT TCACCAGAGG GCTCAACACC GGGATCTATG
251   TTTAAGTTTT AACTCTCGCC TCCAAAGACC ACGATAATTC CTTCCCCAAA
301   GCCCAGCAGC CCCCCAGCCC CGCGCAGCCC CAGCCTGCCT CCCGGCGCCC
351   AGATGCCCGC CATGCCCTCC AGCGGCCCCG GGGACACCAG CAGCTCTGCT
401   GCGGAGCGGG AGGAGGACCG AAAGGACGGA GAGGAGCAGG AGGAGCCGCG
451   TGGCAAGGAG GAGCGCCAAG AGCCCAGCAC CACGGCACGG AAGGTGGGGC
501   GGCCTGGGAG GAAGCGCAAG CACCCCCCGG TGGAAAGCGG TGACACGCCA
551   AAGGACCCTG CGGTGATCTC CAAGTCCCCA TCCATGGCCC AGGACTCAGG
601   CGCCTCAGAG CTATTACCCA ATGGGGACTT GGAGAAGCGG AGTGAGCCCC
651   AGCCAGAGGA GGGGAGCCCT GCTGGGGGGC AGAAGGGCGG GGCCCCAGCA
701   GAGGGAGAGG GTGCAGCTGA GACCCTGCCT GAAGCCTCAA GAGCAGTGGA
751   AAATGGCTGC TGCACCCCCA AGGAGGGCCG AGGAGCCCCT GCAGAAGCGG
801   GCAAAGAACA GAAGGAGACC AACATCGAAT CCATGAAAAT GGAGGGCTCC
851   CGGGGCCGGC TGCGGGGTGG CTTGGGCTGG GAGTCCAGCC TCCGTCAGCG
901   GCCCATGCCG AGGCTCACCT TCCAGGCGGG GGACCCCTAC TACATCAGCA
951   AGCGCAAGCG GGACGAGTGG CTGGCACGCT GGAAAAGGGA GGCTGAGAAG
1001  AAAGCCAAGG TCATTGCAGG AATGAATGCT GTGGAAGAAA ACCAGGGGCC
1051  CGGGGAGTCT CACAAGGTGG AGGAGGCCAG CCCTCCTGCT GTGCAGCAGC
1101  CCACTGACCC CGCATCCCCC ACTGTGGCTA CCACGCCTGA GCCCGTGGGG
1151  TCCGATGCTG GGGACAAGAA TGCCACCAAA GCAGGCGATG ACGAGCCAGA
```

AMENDED SHEET

FIG. 1C-1

1201 GTACGAGGAC GGCCGGGGCT TTGGCATTGG GGAGCTGGTG TGGGGGAAAC

1251 TGCGGGGCTT CTCCTGGTGG CCAGGCCGCA TTGTGTCTTG GTGGATGACG

1301 GGCCGGAGCC GAGCAGCTGA AGGCACCCGC TGGGTCATGT GGTTCGGAGA

1351 CGGCAAATTC TCAGTGGTGT GTGTTGAGAA GCTGATGCCG CTGAGCTCGT

1401 TTTGCAGTGC GTTCCACCAG GCCACGTACA ACAAGCAGCC CATGTACCGC

1451 AAAGCCATCT ACGAGGTCCT GCAGGTGGCC AGCAGCCGCG CGGGGAAGCT

1501 GTTCCCGGTG TGCCACGACA GCGATGAGAG TGACACTGCC AAGGCCGTGG

1551 AGGTGCAGAA CAAGCCCATG ATTGAATGGG CCCTGGGGGG CTTCcAGCAT

1601 TATGGCCCTA AGGGCCTGGA GCCACCAGAA GAAGAGAAGA ATCCCTACAA

1651 AGAAGTGTAC ACGGACATGT GGGTGGAACC TGAGGCAGCT GCATACGCAC

1701 CACCTCCACC AGCCAAAAAG CCCCGGAAGA GCACAGCGGA GAAGCCCAAG

1751 GTCAAGGAGA TTATTGATGA GCGCACAAGA GAGCGGcTGG TGTACGAGGT

1801 GCGGCAGAAG TGCCGGAACA TTGAGGACAT CTGCATCTCC TGTGGGAGCC

1851 TCAATGTTAC CCTGGAACAC CCCCTCTTCG TTGGAGGAAT GTGCCAAAAC

1901 TGCAAGAACT GCTTTCTGGA GTGTGCGTAC CAGTACGACG ACGACGGCTA

1951 CCAGTCCTAC TGCACCATCT GCTGTGGGGG CCGTGAGGTG CTCATGTGCG

2001 GAAACAACAA CTGCTGCAGG TGCTTTTGCG TGGAGTGTGT GGACCTCTTG

2051 GTGGGGCCGG GGGCTGCCCA gGCAGCCATT AAGGAAgACC CCTGGAACTG

2101 CTACATGTGC GGGCACAAGG GTACCTACGG GCTGCTGCGG CGGCGAAAGG

2151 ACTGGCCCTC CCGGCTCCAg ATGTTCTTCG CTAATAACCA CgACCAGgAA

2201 TTTGACCCTC CAAAGGTTTA CCCACCTGTC CCAGCTgAgA AAAGGAAGCC

2251 CATCCGGGTG CTGTCTCTCT TTGATGGAAT CGCTACAGGG CTCCTGGTGC

2301 TGAAGGACTT GGGCATTCAG GTGGACCGCT ACATTGCCTC GGAGGTGTGT

FIG. 1C—2

FORM PCT/RO/134 (JULY 1998)

2351 GAGGACTCCA TCACGGTGGC CATGCTGCGG CACCAGGGGA AGATCATGTA
2401 CGTCGGGGAC GTCCGCAGCG TCACACAGAA GCATATCCAG GAGTGCGGCC
2451 CATTCGATCT GGTGATTGGG GGCAGTCCCT GCAATGACCT CTCCATCGTC
2501 AACCCTGCTC GCAAGGGCCT CTACGAGGGC ACTGGCCGGC TCTTCTTTGA
2551 GTTCTACCGC CTCCTGCATG ATGCGCGGCC CAAGGAGGGA GATGATCGCC
2601 CCTTCTTCTG GCTCTTTGAG AATGTGGTGG CCATGGGCGT TAGTGACAAG
2651 AGGGACATCT CGCGATTTCT CGAGTCCAAC CCTGTGATGA TTGATGCCAA
2701 AGAAGTGTCA GCTGCACACA GGGCCCGCTA CTTCTGGGGT AACCTTCCCG
2751 GTATGAACAG GCCGTTGGCA TCCACTGTGA ATGATAAGCT GGAGCTGCAG
2801 GAGTGTCTGG AGCATGGCAG GATAGCCAAG TTCAGCAAAG TGAGGACCAT
2851 TACTACGAGG TCAAACTCCA TAAAGCAGGG CAAAGACCAG CATTTTCCTG
2901 TCTTCATGAA TGAGAAAGAG GACATCTTAT GGTGCACTGA AATGGAAAGG
2951 GTATTTGGTT TCCCAGTCCA CTATACTGAC GTCTCCAACA TGAGCCGCTT
3001 GGCGAGGCAG AGACTGCTGG GCCGGTCATG GAGCGTGCCA GTCATCCGCC
3051 ACCTCTTCGC TCCGCTGAAG GAGTATTTTG CGTGTGTGTA AGGGACATGG
3101 GGGCAAACTG AGGTAGCGAC ACAAAGTTAA ACAAACAAAC AAAAAACACA
3151 AAACATAATA AAACACCAAG AACATGAGGA TGGAGAGAAG TATCAGCACC
3201 CAGAAGAGAA AAAGGAATTT AAAACAAAAA CCACAGAGGC GGAAATACCG
3251 GAGGGCTTTG CCTTGCGAAA AGGGTTGGAC ATCATCTCCT GATTTTTCAA
3301 TGTTATTCTT CAGTCCTATT TAAAAACAAA ACCAAGCTCC CTTCCCTTCC
3351 TCCCCCTTCC CTTTTTTTTC GGTCAGACCT TTTATTTTCT ACTCTTTTCA
3401 GAGGGGTTTT CTGTTTGTTT GGGTTTTGTT TCTTGCTGTG ACTGAAACAA
3451 GAAGGTTATT GCAGCAAAAA TCAGTAACAA AAAATAGTAA CAATACCTTG
3501 CACAGGAAAG GTGGGAGGAG AGGAAAAAAG GGAAATTTTT AAAGAAATCT

FIG. 1C-3

AMENDED SHEET

09772086 .072301

```
3551 ATATATTGGG TTGTTTTTTT TTTTGTTTTT TGTTTTTTTT TTTTGGGTTT
3601 TTTTTTTTTA CTATATATCT TTTTTTTGTT GTCTCTAGCC TGATCAGATA
3651 GGAGCACAAG CAGGGGACGG AAAGAGAGAG ACACTCAGGC GGCAGCATTC
3701 CCTGCCAGCC ACTGAGCTGT CGTGCCAGCA CCATTCCTGG TCACGCAAAA
3751 CAGAACCCAG TTAGCAGCAG GGAGACGAGA ACACCACACA AGACATTTTT
3801 CTACAGTATT TCAGGTGCCT ACCACACAGG AAACCTTGAA GAAAATCAGT
3851 TTCTAGAAGC CGCTGTTACC TCTTGTTTAC AGTTTATATA TATATGATAG
3901 ATATGAGATA TATATATAAA AGGTACTGTT AACTACTGTA CAACCCGACT
3951 TCATAATGGT GCTTTCAAAC AGCGAGATGA GTAAAAACAT CAGCTTCCAC
4001 GTTGCCTTCT GCGCAAAGGG TTTCACCAAG GATGGAGAAA GGGAGACAGC
4051 TTGCAGATGG CGCGTTCTCA CGGTGGGCTC TTCCCCTTGG TTTGTAACGA
4101 AGTGAAGGAG GAGAACTTGG GAGCCAGGTT CTCCCTGCCA AAAAGGGGGC
4151 TAGATGAGGT GGTCGGGCCC GTGGACAGCT GAGAGTGGGA TTCATCCAGA
4201 CTCATGCAAT AACCCTTTGA TTGTTTTCTA AAAGGAGACT CCCTCGGCAA
4251 GATGGCAGAG CGTACGGAGT CTTCAGGCCC AGTTTCTCAC TTTAGCCAAT
4301 TCGAGGGCTC CTTGTGGTGG GATCAGAACT AATCCAGAGT GTGGGAAAGT
4351 GACAGTCAAA ACCCCACCTG GAGCAAATAA AAAAACATAC AAAACGTAAA
4401 AAAAAAAAAA AAAAAA
```

FIG. 1C-4

AMENDED SHEET

Human DNMT3B1 DNA Sequence:

1 GGCCGCGAAT TCGGCACGAG CCCTGCACGG CCGCCAGCCG GCCTCCCGCC

51 AGCCAGCCCC GACCCGCGGC TCCGCCGCCC AGCCGCGCCC CAGCCAGCCC

101 TGCGGCAGGA AAGCATGAAG GGAGACACCA GGCATCTCAA TGGAGAGGAG

151 GACGCCGGCG GGAGGGAAGA CTCGATCCTC GTCAACGGGG CCTGCAGCGA

201 CCAGTCCTCC GACTCGCCCC CAATCCTGGA GGCTATCCGC ACCCCGGAGA

251 TCAGAGGCCG AAGATCAAGC TCGCGACTCT CCAAGAGGGA GGTGTCCAGT

301 CTGCTAAGCT ACACACAGGA CTTGACAGGC GATGGCGACG GGGAAGATGG

351 GGATGGCTCT GACACCCCAG TCATGCCAAA GCTCTTCCGG GAAACCAGGA

401 CTCGTTCAGA AAGCCCAGCT GTCCGAACTC GAAATAACAA CAGTGTCTCC

451 AGCCCGGAGA GGCACAGGCC TTCCCCACGT TCCACCCGAG GCCGGCAGGG

501 CCGCAACCAT GTGGACGAGT CCCCCGTGGA GTTCCCGGCT ACCAGGTCCC

551 TGAGACGGCG GGCAACAGCA TCGGCAGGAA CGCCATGGCC GTCCCCTCCC

601 AGCTCTTACC TTACCATCGA CCTCACAGAC GACACAGAGG ACACACATGG

651 GACGCCCCAG AGCAGCAGTA CCCCCTACGC CCGCCTAGCC CAGGACAGCC

701 AGCAGGGGGG CATGGAGTCC CCGCAGGTGG AGGCAGACAG TGGAGATGGA

751 GACAGTTCAG AGTATCAGGA TGGGAAGGAG TTTGGAATAG GGGACCTCGT

801 GTGGGGAAAG ATCAAGGGCT TCTCCTGGTG GCCCGCCATG GTGGTGTCTT

851 GGAAGGCCAC CTCCAAGCGA CAGGCTATGT CTGGCATGCG GTGGGTCCAG

901 TGGTTTGGCG ATGGCAAGTT CTCCGAGGTC TCTGCAGACA AACTGGTGGC

951 ACTGGGGCTG TTCAGCCAGC ACTTTAATTT GGCCACCTTC AATAAGCTCG

1001 TCTCCTATCG AAAAGCCATG TACCATGCTC TGGAGAAAGC TAGGGTGCGA

1051 GCTGGCAAGA CCTTCCCCAG CAGCCCTGGA GACTCATTGG AGGACCAGCT

1101 GAAGCCCATG TTGGAGTGGG CCCACGGGGG CTTCAAGCCC ACTGGGATCG

1151 AGGGCCTCAA ACCCAACAAC ACGCAACCAG TGGTTAATAA GTCGAAGGTG

FIG. 1D−1

PCT/US99/14373 09.06.2000

1201 CGTCGTGCAG GCAGTAGGAA ATTAGAATCA AGGAAATACG AGAACAAGAC

1251 TCGAAGACGC ACAGCTGACG ACTCAGCCAC CTCTGACTAC TGCCCCGCAC

1301 CCAAGCGCCT CAAGACAAAT TGCTATAACA ACGGCAAAGA CCGAGGGGAT

1351 GAAGATCAGA GCCGAGAACA AATGGCTTCA GATGTTGCCA ACAACAAGAG

1401 CAGCCTGGAA GATGGCTGTT TGTCTTGTGG CAGGAAAAAC CCCGTGTCCT

1451 TCCACCCTCT CTTTGAGGGG GGGCTCTGTC AGACATGCCG GGATCGCTTC

1501 CTTGAGCTGT TTTACATGTA TGATGACGAT GGCTATCAGT CTTACTGCAC

1551 TGTGTGCTGC GAGGGCCGAG AGCTGCTGCT TTGCAGCAAC ACGAGCTGCT

1601 GCCGGTGTTT CTGTGTGGAG TGCCTGGAGG TGCTGGTGGG CACAGGCACA

1651 GCGGCCGAGG CCAAGCTTCA GGAGCCCTGG AGCTGCTACA TGTGTCTCCC

1701 GCAGCGCTGT CATGGCGTCC TGCGGCGCCG GAAGGACTGG AACGTGCGCC

1751 TGCAGGCCTT CTTCACCAGT GACACGGGGC TTGAATACGA AGCCCCCAAG

1801 CTGTACCCTG CCATTCCCGC AGCCCGAAGG CGGCCCATTC GAGTCCTGTC

1851 ATTGTTTGAT GGCATCGCGA CAGGCTACCT AGTCCTCAAA GAGTTGGGCA

1901 TAAAGGTAGG AAAGTACGTC GCTTCTGAAG TGTGTGAGGA GTCCATTGCT

1951 GTTGGAACCG TGAAGCACGA GGGGAATATC AAATACGTGA ACGACGTGAG

2001 GAACATCACA AAGAAAAATA TTGAAGAATG GGGCCCATTT GACTTGGTGA

2051 TTGGCGGAAG CCCATGCAAC GATCTCTCAA ATGTGAATCC AGCCAGGAAA

2101 GGCCTGTATG AGGGTACAGG CCGGCTCTTC TTCGAATTTT ACCACCTGCT

2151 GAATTACTCA CGCCCCAAGG AGGGTGATGA CCGGCCGTTC TTCTGGATGT

2201 TTGAGAATGT TGTAGCCATG AAGGTTGGCG ACAAGAGGGA CATCTCACGG

2251 TTCCTGCAGT GTAATCCAGT GATGATTGAT GCCATCAAAG TTTCTGCTGC

2301 TCACAGGGCC CGATACTTCT GGGGCAACCT ACCCGGGATG AACAGGCCCG

2351 TGATAGCATC AAAGAATGAT AAACTCGAGC TGCAGGACTG CTTGGAATAC

2401 AATAGGATAG CCAAGTTAAA GAAAGTACAG ACAATAACCA CCAAGTCGAA

FIG. 1D-2

PCT/US99/14373 09-06-2000

2451 CTCGATCAAA CAGGGGAAAA ACCAACTTTT CCCTGTTGTC ATGAATGGCA
2501 AAGAAGATGT TTTGTGGTGC ACTGAGCTCG AAAGGATCTT TGGCTTTCCT
2551 GTGCACTACA CAGACGTGTC CAACATGGGC CGTGGTGCCC GCCAGAAGCT
2601 GCTGGGAAGG TCCTGGAGCG TGCCTGTCAT CCGACACCTC TTCGCCCCTC
2651 TGAAGGACTA CTTTGCATGT GAATAGTTCC AGCCAGGCCC CAAGCCCACT
2701 GGGGTGTGTG GCAGAGCCAG GACCCAGGAG GTGTGATTCC TGAAGGCATC
2751 CCCAGGCCCT GCTCTTCCTC AGCTGTGTGG GTCATACCGT GTACCTCAGT
2801 TCCCTCTTGC TCAGTGGGGG CAGAGCCACC TGACTCTTGC AGGGGTAGCC
2851 TGAGGTGCCG CCTCCTTGTG CACAAATCAG ACCTGGCTGC TTGGAGCAGC
2901 CTAACACGGT GCTCATTTTT TCTTCTCCTA AAACTTTAAA ACTTGAAGTA
2951 GGTAGCAACG TGGCTTTTTT TTTTTCCCTT CCTGGGTCTA CCACTCAGAG
3001 AAACAATGGC TAAGATACCA AAACCACAGT GCCGACAGCT CTCCAATACT
3051 CAGGTTAATG CTGAAAAATC ATCCAAGACA GTTATTGCAA GAGTTTAATT
3101 TTTGAAAACT GGGTACTGCT ATGTGTTTAC AGACGTGTGC AGTTGTAGGC
3151 ATGTAGCTAC AGGACATTTT TAAGGGCCCA GGATCGTTTT TTCCCAGGGC
3201 AAGCAGAAGA GAAAATGTTG TATATGTCTT TTACCCGGCA CATTCCCCTT
3251 GCCTAAATAC AAGGGCTGGA GTCTGCACGG GACCTATTAG AGTATTTTCC
3301 ACAATGATGA TGATTTCAGC AGGGATGACG TCATCATCAC ATTCAGGGCT
3351 ATTTTTTCCC CCACAAACCC AAGGGCAGGG GCCACTCTTA GCTAAATCCC
3401 TCCCCGTGAC TGCAATAGAA CCCTCTGGGG AGCTCAGGAA GGGGTGTGCT
3451 GAGTTCTATA ATATAAGCTG CCATATATTT TGTAGACAAG TATGGCTCCT
3501 CCATATCTCC CTCTTCCCTA GGAGAGGAGT GTGAAGCAAC GAGCTTAGAT
3551 AAGACACCCC CTCAAACCCA TTCCCTCTCC AGGAGACCTA CCCTCCACAG
3601 GCACAGGTCC CCAGATGAGA AGTCTGCTAC CCTCATTTCT CATCTTTTTA
3651 CTAAACTCAG AGGCAGTGAC AGCAGTCAGG GACAGACATA CATTTCTCAT

FIG. 1D-3

AMENDED SHEET

09720086 .071301

TOET40 980627 60

```
3701  ACCTTCCCCA CATCTGAGAG ATGACAGGGA AAACTGCAAA GCTCGGTGCT
3751  CCCTTTGGAG ATTTTTTAAT CCTTTTTTAT TCCATAAGAA GTCGTTTTTA
3801  GGGAGAACGG GAATTCAGAC AAGCTGCATT TCAGAAATGC TGTCATAATG
3851  GTTTTTAACA CCTTTTACTC TTCTTACTGG TGCTATTTTG TAGAATAAGG
3901  AACAACGTTG ACAAGTTTTG TGGGGCTTTT TATACACTTT TTAAAATCTC
3951  AAACTTCTAT TTTTATGTTT AACGTTTTCA TTAAAATTTT TTTGTAACTG
4001  GAGCCACGAC GTAACAAATA TGGGGAAAAA ACTGTGCCTT GTTTCAACAG
4051  TTTTTGCTAA TTTTTAGGCT GAAAGATGAC GGATGCCTAG AGTTTACCTT
4101  ATGTTTAATT AAAATCAGTA TTTGTCTAAA AAAAAAAAAA AAAAA
```

FIG. 1D-4

AMENDED SHEET

Mouse Dnmt3a Protein

1 MPSSGPGDTS SSSLEREDDR KEGEEQEENR GKEERQEPSA TARKVGRPGR

51 KRKHPPVESS DTPKDPAVTT KSQPMAQDSG PSDLLPNGDL EKRSEPQPEE

101 GSPAAGQKGG APAEGEGTET PPEASRAVEN GCCVTKEGRG ASAGEGKEQK

151 QTNIESMKME GSRGRLRGGL GWESSLRQRP MPRLTFQAGD PYYISKRKRD

201 EWLARWKRDA EKKAKVIAVM NAVEENQASG ESQKVEEASP PAVQQPTDPA

251 SPTVATTPEP VGGDAGDKNA TKAPDDEPEY EDGRGFGIGE LVWGKLRGFS

301 WWPGRIVSWW MTGRSRAAEG TRWVMWFGDG KFSVVCVEKL MPLSSFCSAF

351 HQATYNKQPM YRKAIYEVLQ VASSRAGKLF PACHDSDESD SGKAVEVQNK

401 QMIEWALGGF QPSGPKGLEP PEEEKNPYKE VYTDMWVEPE AAAYAPPPPA

451 KKPRKSTTEK PKVKEIIDER TRERLVYEVR QKCRNIEDIC ISCGSLNVTL

501 EHPFFIGGMC QNCKNCFLEC AYQYDDDGYQ SYCTICCGGR EVLMCGNNNC

551 CRCFCVECVD LLVGPGAAQA AIKEDPWNCY MCGHKGTYGL LRRREDWPSR

601 LQMFFANNHD QEFDPPKVYP PVPAEKRKPI RVLSLFDGIA TGLLVLKDLG

651 IQVDRYIASE VCEDSITVGM VRHQGKIMYV GDVRSVTQKH IQEWGPFDLV

701 IGGSPCNDLS IVNPARKGLY EGTGRLFFEF YRLLHDARPK EGDDRPFFWL

751 FENVVAMGVS DKRDISRFLE SNPVMIDAKE VSAAHRARYF WGNLPGMNRP

801 LASTVNDKLE LQECLEHGRI AKFSKVRTIT TRSNSIKQGK DQHFPVFMNE

851 KEDILWCTEM ERVFGFPVHY TDVSNMSRLA RQRLLGRSWS VPVIRHLFAP

901 LKEYFACV*

FIG. 2A

AMENDED SHEET

09772086 . 070301

Mouse Dnmt3b1 Protein

1 MKGDSRHLNE EEGASGYEEC IIVNGNFSDQ SSDTKDAPSP PVLEAICTEP

51 VCTPETRGRR SSSRLSKREV SSLLNYTQDM TGDGDRDDEV DDGNGSDILM

101 PKLTRETKDT RTRSESPAVR TRHSNGTSSL ERQRASPRIT RGRQGRHHVQ

151 EYPVEFPATR SRRRRASSSA STPWSSPASV DFMEEVTPKS VSTPSVDLSQ

201 DGDQEGMDTT QVDAESIYGD STEYQDDKEF GIGDLVWGKI KGFSWWPAMV

251 VSWKATSKRQ AMPGMRWVQW FGDGKFSEIS ADKLVALGLF SQHFNLATFN

301 KLVSYRKAMY HTLEKARVRA GKTFSSSPGE SLEDQLKPML EWAHGGFKPT

351 GIEGLKPNKK QPVVNKSKVR RSDSRNLEPR RRENKSRRRT TNDSAASESP

401 PPKRLKTNSY GGKDRGEDEE SRERMASEVT NNKGNLEDRC LSCGKKNPVS

451 FHPLFEGGLC QSCRDRFLEL FYMYDEDGYQ SYCTVCCEGR ELLLCSNTSC

501 CRCFCVECLE VLVGAGTAED AKLQEPWSCY MCLPQRCHGV LRRRKDWNMR

551 LQDFFTTDPD LEEFEPPKLY PAIPAAKRRP IRVLSLFDGI ATGYLVLKEL

601 GIKVEKYIAS EVCAESIAVG TVKHEGQIKY VNDVRKITKK NIEEWGPFDL

651 VIGGSPCNDL SNVNPARKGL YEGTGRLFFE FYHLLNYTRP KEGDNRPFFW

701 MFENVVAMKV NDKKDISRFL ACNPVMIDAI KVSAAHRARY FWGNLPGMNR

751 PVMASKNDKL ELQDCLEFSR TAKLKKVQTI TTKSNSIRQG KNQLFPVVMN

801 GKDDVLWCTE LERIFGFPAH YTDVSNMGRG ARQKLLGRSW SVPVIRHLFA

851 PLKDYFACE*

FIG. 2B

AMENDED SHEET

Human DNMT3A Protein

```
  1 MPAMPSSGPG DTSSSAAERE EDRKDGEEQE EPRGKEERQE PSTTARKVGR
 51 PGRKRKHPPV ESGDTPKDPA VISKSPSMAQ DSGASELLPN GDLEKRSEPQ
101 PEEGSPAGGQ KGGAPAEGEG AAETLPEASR AVENGCCTPK EGRGAPAEAG
151 KEQKETNIES MKMEGSRGRL RGGLGWESSL RQRPMPRLTF QAGDPYYISK
201 RKRDEWLARW KREAEKKAKV IAGMNAVEEN QGPGESHKVE EASPPAVQQP
251 TDPASPTVAT TPEPVGSDAG DKNATKAGDD EPEYEDGRGF GIGELVWGKL
301 RGFSWWPGRI VSWWMTGRSR AAEGTRWVMW FGDGKFSVVC VEKLMPLSSF
351 CSAFHQATYN KQPMYRKAIY EVLQVASSRA GKLFPVCHDS DESDTAKAVE
401 VQNKPMIEWA LGGFQHYGPK GLEPPEEEKN PYKEVYTDMW VEPEAAAYAP
451 PPPAKKPRKS TAEKPKVKEI IDERTRERLV YEVRQKCRNI EDICISCGSL
501 NVTLEHPLFV GGMCQNCKNC FLECAYQYDD DGYQSYCTIC CGGREVLMCG
551 NNNCCRCFCV ECVDLLVGPG AAQAAIKEDP WNCYMCGHKG TYGLLRRRKD
601 WPSRLQMFFA NNHDQEFDPP KVYPPVPAEK RKPIRVLSLF DGIATGLLVL
651 KDLGIQVDRY IASEVCEDSI TVGMVRHQGK IMYVGDVRSV TQKHIQEWGP
701 FDLVIGGSPC NDLSIVNPAR KGLYEGTGRL FFEFYRLLHD ARPKEGDDRP
751 FFWLFENVVA MGVSDKRDIS RFLESNPVMI DAKEVSAAHR ARYFWGNLPG
801 MNRPLASTVN DKLELQECLE HGRIAKFSKV RTITTRSNSI KQGKDQHFPV
851 FMNEKEDILW CTEMERVFGF PVHYTDVSNM SRLARQRLLG RSWSVPVIRH
901 LFAPLKEYFA CV*
```

FIG. 2C

AMENDED SHEET

FOCT/US99/14373

Human DNMT3B1 Prolein

1 MKGDTRHLNG EEDAGGREDS ILVNGACSDQ SSDSPPILEA IRTPEIRGRR

51 SSSRLSKREV SSLLSYTQDL TGDGDGEDGD GSDTPVMPKL FRETRTRSES

101 PAVRTRNNNS VSSRERHRPS PRSTRGRQGR NHVDESPVEF PATRSLRRRA

151 TASAGTPMPS PPSSYLTIDL TDDTEDTHGT PQSSSTPYAR LAQDSQQGGM

201 ESPQVEADSG DGDSSEYQDG KEFGIGDLVW GKIKGFSWWP AMVVSWKATS

251 KRQAMSGMRW VQWFGDGKFS EVSADKLVAL GLFSQHFNLA TFNKLVSYRK

301 AMYHALEKAR VRAGKTFPSS PGDSLEDQLK PMLEWAHGGF KPTGIEGLKP

351 NNTQPVVNKS KVRRAGSRKL ESRKYENKTR RRTADDSATS DYCPAPKRLK

401 TNCYNNGKDR GDEDQSREQM ASDVANNKSS LEDGCLSCGR KNPVSFHPLF

451 EGGLCQTCRD RFLELFYMYD DDGYQSYCTV CCEGRELLLC SNTSCCRCFC

501 VECLEVLVGT GTAAEAKLQE PWSCYMCLPQ RCHGVLRRRK DWNVRLQAFF

551 TSDTGLEYEA PKLYPAIPAA RRRPIRVLSL FDGIATGYLV LKELGIKVGK

601 YVASEVCEES IAVGTVKHEG NIKYVNDVRN ITKKNIEEWG PFDLVIGGSP

651 CNDLSNVNPA RKGLYEGTGR LFFEFYHLLN YSRPKEGDDR PFFWMFENVV

701 AMKVGDKRDI SRFLECNPVM IDAIKVSAAH RARYFWGNLP GMNRPVIASK

751 NDKLELQDCL EYNRIAKLKK VQTITTKSNS IKQGKNQLFP VVMNGKEDVL

801 WCTELERIFG FPVHYTDVSN MGRGARQKLL GRSWSVPVIR HLFAPLKDYF

851 ACE*

FIG. 2D

AMENDED SHEET

```
Dnmt3a    1 MPSSGPGDTSSSSLEREDDRKEGEEQEENRGKEERQEPSATARKVGRPGR 50
Dnmt3a   51 KRKHPPVESSDTPKDPAVTTKSQPMAQDSGPSD....LLPNGDLEKRSEP 96
Dnmt3b    1 ..................MKGDSRHLNEEEGASGYEECIIVNGNFSDQSSD 33
Dnmt3a   97 QPEEGSP....AAGQKGGAPAEGEGTETPPEAS.RAVENGCCVTKE..GR 139
Dnmt3b   34 TKDAPSPPVLEAICTEPVCTPETRGRRSSSRLSKREVSSLLNYTQDMTGD 83
Dnmt3a  140 G.....ASAGEG......KEQKQTNIESMKMEGSRGRLRGGLGWESSLRQ 178
Dnmt3b   84 GDRDDEVDDGNGSDILMPKLTRETKDTRTRSESPAVRTRHSNGTSSLERQ 133
Dnmt3a  179 RPMPRLTFQAGDPYYISKRKRDEWLARWKRDAEKKAKVIAVMNAVEENQA 228
Dnmt3b  134 RASPRITRGRQGRHHV.....QEYPVEFPATRSRRRRASSSASTPWSSPA 178
Dnmt3a  229 SGESQKVEEASPPAVQQPTDPASPTVATTPEPVGGDAGDKNATKAPDDEP 278
Dnmt3b  179 SVDF..MEEVTPKSVSTP....SVDLSQDGDQEGMDTTQVDAESIYGDST 222
Dnmt3a  279 EYEDGRGFGIGELVWGKLRGFSWWPGRIVSWWMTGRSRAAEGTRWVMWFG 328
Dnmt3b  223 EYQDDKEFGIGDLVWGKIKGFSWWPAMVVSWKATSKRQAMPGMRWVQWFG 272
Dnmt3a  329 DGKFSVVCVEKLMPLSSFCSAFHQATYNKQPMYRKAIYEVLQVASSRAGK 378
Dnmt3b  273 DGKFSEISADKLVALGLFSQHFNLATFNKLVSYRKAMYHTLEKARVRAGK 322
Dnmt3a  379 LFPACHDSDESDSGKAVEVQNKQMIEWALGGFQPSGPKGLEPPEEEK..N 426
Dnmt3b  323 TF.......SSSPGESLEDQLKPMLEWAHGGFKPTGIEGLKPNKKQPVVN 365
Dnmt3a  427 PYKEVYTDMW.VEP.............EAAAYAPPPPAKKPRKSTTEKPK 462
Dnmt3b  366 KSKVRRSDSRNLEPRRRENKSRRRTTNDSAASESPPPKRLKTNSYGGKDR 415
```

FIG.3A-1

AMENDED SHEET

```
Dnmt3a  463 VKEIIDERTRERLVYEVRQKCRNIEDICISCGSLNVTLEHPFFIGGMCQN 512
Dnmt3b  416 GE...DEESRERMASEVTNNKGNLEDRCLSCGKKNPVSFHPLFEGGLCQS 462

Dnmt3a  513 CKNCFLECAYQYDDDGYQSYCTICCGGREVLMCGNNNCCRCFCVECVDLL 562
Dnmt3b  463 CRDRFLELFYNYDEDGYQSYCTVCCEGRELLLCSNTSCCRCFCVECLEVL 512

Dnmt3a  563 VGPGAAQAAIKEDPWNCYMCGHKGTYGLLRRREDWPSRLQMFFANNHD.Q 611
Dnmt3b  513 VGAGTAEDAKLQEPWSCYMCLPQRCHGVLRRRKDWNMRLQDFFTTDPDLE 562

Dnmt3a  612 EFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIASEV 661
Dnmt3b  563 EFEPPKLYPAIPAAKRRPIRVLSLFDGIATGYLVLKELGIKVEKYIASEV 612

Dnmt3a  662 CEDSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSI 711
Dnmt3b  613 CAESIAVGTVKHEGQIKYVNDVRKITKKNIEEWGPFDLVIGGSPCNDLSN 662

Dnmt3a  712 VNPARKGLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSD 761
Dnmt3b  663 VNPARKGLYEGTGRLFFEFYHLLNYTRPKEGDNRPFFWMFENVVAMKVND 712

Dnmt3a  762 KRDISRFLESNPVMIDAKEVSAAHRARYFWGNLPGMNRPLASTVNDKLEL 811
Dnmt3b  713 KKDISRFLACNPVMIDAIKVSAAHRARYFWGNLPGMNRPVMASKNDKLEL 762

Dnmt3a  812 QECLEHGRIAKFSKVRTITTRSNSIKQGKDQHFPVFMNEKEDILWCTEME 861
Dnmt3b  763 QDCLEFSRTAKLKKVQTITTKSNSIRQGKNQLFPVVMNGKDDVLWCTELE 812

Dnmt3a  862 RVFGFPVHYTDVSNMSRLARQRLLGRSWSVPVIRHLFAPLKEYFACV*   909
Dnmt3b  813 RIFGFPAHYTDVSNMGRGARQKLLGRSWSVPVIRHLFAPLKDYFACE*   860
```

FIG.3A-2

AMENDED SHEET

```
DNMT3A    1 MPAMPSSGPGDTSSSAAEREEDRKDGEEQEEPRGKEERQEPSTTARKVGR

DNMT3A   51 PGRKRKHPPVESGDTPKDPAVISKSPSMAQDSGASELLPNGDLEKRSEPQ
                                |              |     | |.     :. |:
DNMT3B    1 ...............MKGDTRHLNGEEDAGGREDSILVNGACSDQSSDSP

DNMT3A  101 PEEGSPAGGQKGGAPAEGEGAAETLPEASRAVENGCCTPKEGRGAPAEAG
              |   .    :  |  |    .   .          :.          |  .:
DNMT3B   36 PILEAIRTPEIRGGWASSRLSKREVSSLLSYTQDLTGDGDGEDGDGSDTP

DNMT3A  151 KEQKETNIESMKMEGSRGRLRGGLGWESSLRQRPMPRLTFQAGDPYYISK
                 |       : |    | |       |  | || || |       :: .
DNMT3B   86 VMPKLFRETRTRSESPAVRTRNNNSVSSRERHRPSPRSTRGRQGRNHVDE

DNMT3A  201 RKRDEWLARWKREAEKKAKVIAGMNAVEENQGPGESHKVEEASPPAVQQP
              :    |  |    .   .          .   |:      |
DNMT3B  136 SPVEFPATRSLRRRATASAGTPWPSPPSSYLTIDLTDDTEDTH..GTPQS

DNMT3A  251 TDPASPTVATTPEPVGSDAGDKNATKAGDDEPEYEDGRGFGIGELVWGKL
              .     .|   :  | :.    |     |  ||:||: ||||:|||||:
DNMT3B  184 SSTPYARLAQDSQQGGMESPQVEADSGDGDSSEYQDGKEFGIGDLVWGKI

DNMT3A  301 RGFSWWPGRIVSWWMTGRSRAAEGTRWVMWFGDGKFSVVCVEKLMPLSSF
             :||||||  :|||  | : .|  | ||| ||||||||  |  :||. |  |
DNMT3B  234 KGFSWWPAMVVSWKATSKRQAMSGMRWVQWFGDGKFSEVSADKLVALGLF

DNMT3A  351 CSAFHQATYNKQPMYRKAIYEVLQVASSRAGKLFPVCHDSDESDTAKAVE
              |. ||:||   ||||.|  |: |  |||| ||       |    ..|
DNMT3B  284 SQHFNLATFNKLVSYRKAMYHALEKARVRAGKTFP.......SSPGDSLE

DNMT3A  401 VQNKPMIEWALGGFQHYGPKGLEP....PEEEKNPYKEVYTDMWVE.. .
             | |||:||| |||.  | .||.|    |   |.  :   .
DNMT3B  327 DQLKPMLEWAHGGFKPTGIEGLKPNNTQPVVNKSKVRRAGSRKLESRKYE

DNMT3A  443 .......PEAAAYAPPPPAKKPRKSTAEKPKVKEIIDERTRERLVYEVRQ
                  : .|  .   || |  |.          ::..||.:  :|
DNMT3B  377 NKTRRRTADDSATSDYCPAPKRLKTNCYNNGKDRGDEDQSREQMASDVAN
```

FIG.3B-1

AMENDED SHEET

```
DNMT3A  486 KCRNIEDICISCGSLNVTLEHPLFVGGMCQNCKNCFLECAYQYDDDGYQS
DNMT3B  427 NKSSLEDGCLSCGRKNPVSFHPLFEGGLCQTCRDRFLELFYMYDDDGYQS

DNMT3A  536 YCTICCGGREVLMCGNNNCCRCFCVECVDLLVGPGAAQAAIKEDPWNCYM
DNMT3B  477 YCTVCCEGRELLLCSNTSCCRCFCVECLEVLVGTGTAAEAKLQEPWSCYM

DNMT3A  586 CGHKGTYGLLRRRKDWPSRLQMFFANNHDQEFDPPKVYPPVPAEKRKPIR
DNMT3B  527 CLPQRCHGVLRRRKDWNVRLQAFFTSDTGLEYEAPKLYPAIPAARRRPIR

DNMT3A  636 VLSLFDGIATGLLVLKDLGIQVDRYIASEVCEDSITVGMVRHQGKIMYVG
DNMT3B  577 VLSLFDGIATGYLVLKELGIKVGKYVASEVCEESIAVGTVKHEGNIKYVN

DNMT3A  686 DVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARKGLYEGTGRLFFEFY
DNMT3B  627 DVRNITKKNIEEWGPFDLVIGGSPCNDLSNVNPARKGLYEGTGRLFFEFY

DNMT3A  736 RLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLESNPVMIDAKEV
DNMT3B  677 HLLNYSRPKEGDDRPFFWMFENVVAMKVGDKRDISRFLECNPVMIDAIKV

DNMT3A  786 SAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITT
DNMT3B  727 SAAHRARYFWGNLPGMNRPVIASKNDKLELQDCLEYNRIAKLKKVQTITT

DNMT3A  836 RSNSIKQGKDQHFPVFMNEKEDILWCTEMERVFGFPVHYTDVSNMSRLAR
DNMT3B  777 KSNSIKQGKNQLFPVVMNGKEDVLWCTELERIFGFPVHYTDVSNMGRGAR

DNMT3A  886 QRLLGRSWSVPVIRHLFAPLKEYFACV*
DNMT3B  827 QKLLGRSWSVPVIRHLFAPLKDYFACE*
```

FIG.3B-2

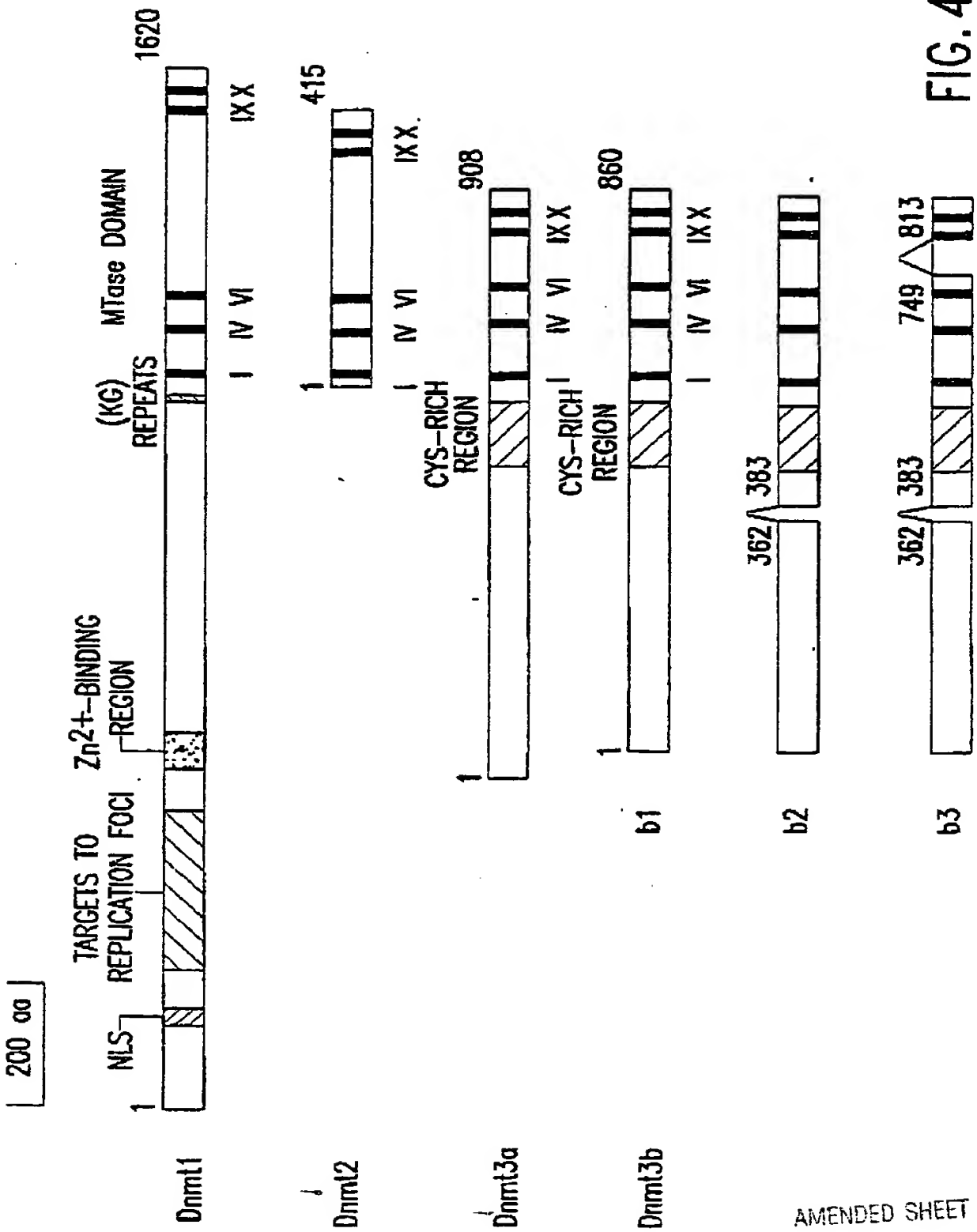
200 aa
NLS
TARGETS TO REPLICATION FOCI
Zn2+-BINDING REGION
(KG) REPEATS
MTase DOMAIN
1620
415
908
860
CYS-RICH REGION
362
383
749
813
I IV VI IX X
Dnmt1
Dnmt2
Dnmt3a
Dnmt3b
b1
b2
b3

FIG. 4A

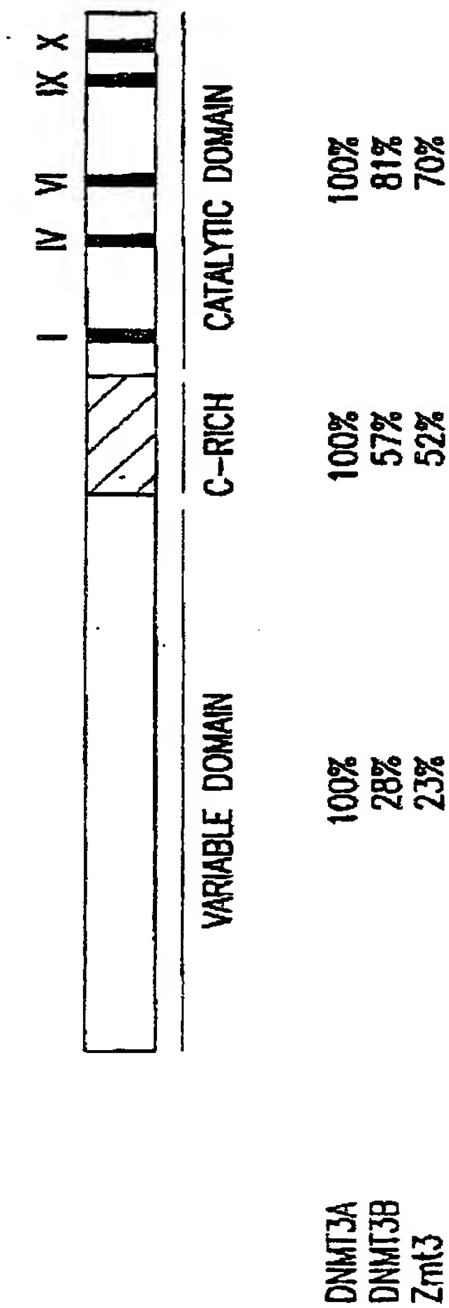

FIG. 4B

AMENDED SHEET

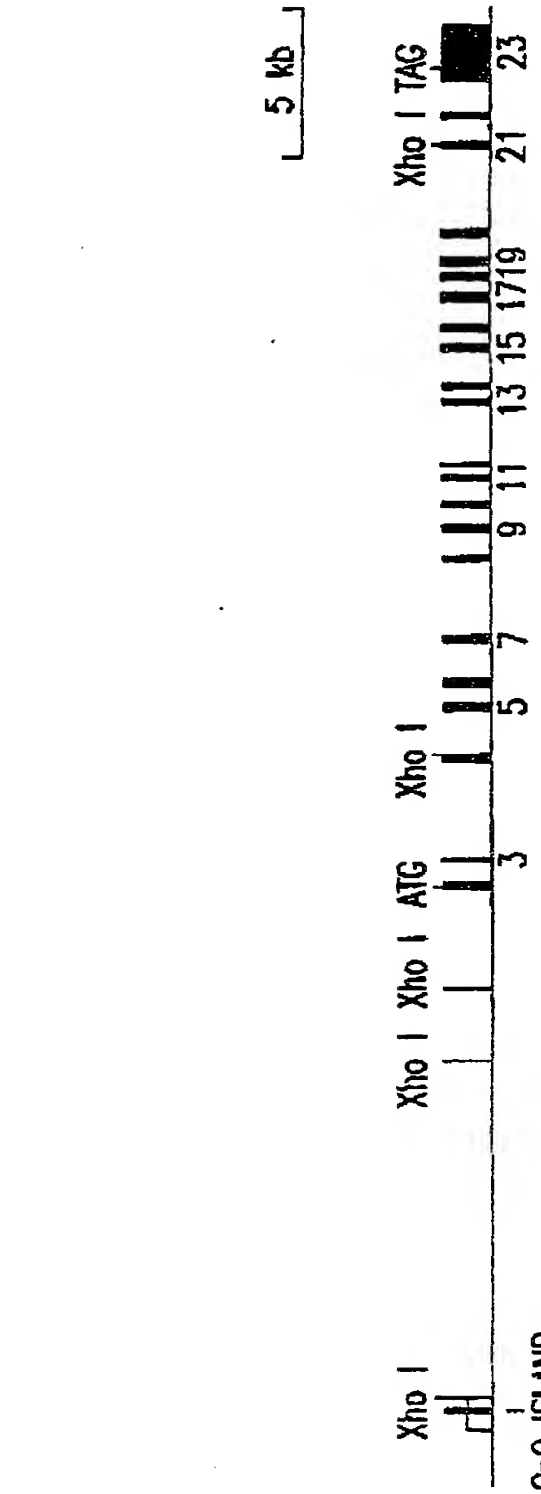

FIG. 4C

AMENDED SHEET

T06L70 "980024/60

```
Exon1  (>=90bp) CGGCAGgtgagcgccccgggg.intron(17618bp).tggcttctcccacagGAAAGC
Exon2  ( 148bp) TCAGAGgtggctgggcagtgg.intron(  887bp).CTGTTTCCTCTACAGGCCGAA
Exon3  (  62bp) ACACAGgtatggtctctgctc.intron( 3343bp).tgtttccttataaagGACTTG
Exon4  ( 102bp) CCAGCTgtaagtagccacacc.intron( 1642bp).ctctcttgcttctagGTCCGA
Exon5  ( 125bp) ACCAGGgttgttccccagatg.intron(  602bp).tccttctgtccacagTCCCTG
Exon6  ( 222bp) TATCAGgtatggccgagaggg.intron( 1403bp).tgggttttcttccagGATGGG
Exon7  ( 159bp) TCCGAGgtgagtccggggaag.intron( 2588bp).gtctttctcttttagGTCTCT
Exon8  ( 108bp) CTGGAGgtaacatgggatgag.intron(  917bp).actctgcctttgcagAAAGCT
Exon9  ( 145bp) AACCAGgtgggaatgagtccc.intron(  765bp).ttttcccctcaaaagTGGTTA
Exon10 (  60bp) AATACGgtatttccttcctgt.intron( 1813bp).aattacctttcacagAGAACA
Exon11 ( 126bp) GCCGAGgtgattgttgggtac.intron(  115bp).ttcttttctcaatagAACAAA
Exon12 (  45bp) TGGAAGgtaacgttctctccc.intron( 1095bp).ctgtttttcttacagATGGCT
Exon13 (  80bp) TGCCGGgtaagtcctcctact.intron(  417bp).ctctctggctgccagGATCGC
Exon14 ( 113bp) CTGCCGgtgagcactgggccc.intron( 1160bp).tgccactgggtccagGTGTTT
Exon15 ( 184bp) GAATACgtaagccacaggctc.intron(  600bp).ttccttacctggcagGAAGCC
Exon16 (  85bp) CGACAGgtgagttcggggaac.intron(  824bp).ctctggcccccacagGCTACC
Exon17 ( 146bp) AAAAATgtgagggcagtctgt.intron(  536bp).gtctctctctttcagATTGAA
Exon18 (  91bp) TGTATGgtgagcatccttctc.intron(  352bp).cttttctgagcacagAGGGTA
Exon19 ( 149bp) CTGGAGgtgagggaatctggg.intron(  958bp).tctttctccccacagTGTAAT
Exon20 (  86bp) GAACAGgtaacaaagggctct.intron( 2867bp).tttggctgttcccagGCCCGT
Exon21 (  70bp) GCCAAGttaaagaaagtacag.intron(  801bp).catttgttctccagTTAAAG
Exon22 ( 119bp) CGAAAGgtgagcaaggctgca.intron( 1434bp).ctccggtacccccagGATCTT
Exon23 (1585bp)
```

FIG.4D

AMENDED SHEET

	I	IV	VI
DNMT1	DVFSGCGGLSEGFHQAG	DVEMLCGGPPCQGFSGMNR	YRPRFFLLENVRNFVSFKR
Dnmt1	DVFSGCGGLSEGFHQAG	DVEMLCGGPPCQGFSGMNR	YRPRFFLLENVRNFVSYRR
MET1(Ath)	DIFAGCGGLSHGLKKAG	QVDFINGGPPCQGFSGMNR	FRPRYFLLENVRTFVSFNK
Masc1	DTFCGGGGVSLGARQAG	HVDILHLSPPCQTFSRAHT	VRPRLFTVEETDGIMDRQS
Masc2	DIFAGCGGLTLGLDLSG	EVDFIYGGPPCQGFSGVNR	YKPRFVLLENVKGLITTKL
Dnmt2	ELYSGIGGMHHALRESH	SFNMILMSPPCQPFTRIGL	KLPKYILLENVKGFEVSST
M.Spr	SLFSGIGAFEAALRNIG	EFDLLVGGSPCQSFSVAGH	KQPKFFVFENVKGLINHDK
DNMT3A	SLFDGIATGLLVLKDLG	PFDLVIGGSPCNDLSIVNP	DRPFFWLFENVVAMGVSDK
Dnmt3a	SLFDGIATGLLVLKDLG	PFDLVIGGSPCNDLSIVNP	DRPFFWLFENVVAMGVSDK
DNMT3B	SLFDGIATGYLVLKELG	PFDLVIGGSPCNDLSNVNP	DRPFFWMFENVVAMKVGDK
Dnmt3b	SLFDGIATGYLVLKELG	PFDLVIGGSPCNDLSNVNP	NRPFFWMFENVVAMKVNDK
Zmt3	SLFDGIATGYLVLRDLG	PFDLLIGGSPCNDLSIVNP	PQPFFWLFENVTFMQTHVK
consensus	--F-G-----------G	------GG-PC---S--N-	--P-F---ENV--------

	IX	X
DNMT1	RVVSVRECARSQGFP	LFGNILDKHRQVGNAVPPPLAKAIG
Dnmt1	RVVSVRECARSQGFP	FFGNILDRHRQVGNAVPPPLAKAIG
MET1(Ath)	RILTVRECARSQGFP	FAGNINHKHRQIGNAVPPPLAFALG
Masc1	RKFTVRELACIQGFP	FVGTLTDKRRIIGNAVPPPLSAAIM
Masc2	RVYTVRELARAQGFP	GLGGVKKWHRNIGNAVPVPLGEQIG
Dnmt2	RYFTPKEIANLQGFP	EKTTVKQRYRLLGNSLNVHVVAKLL
M.Spr	RRLTPLECFRLQAFD	AGISNSQLYKQTGNSITVTVLESIF
DNMT3A	DILWCTEMERVFGFP	SNMSRLARQRLLGRSWSVPVIRHLF
Dnmt3a	DILWCTEMERVFGFP	SNMSRLARQRLLGRSWSVPVIRHLF
DNMT3B	DVLWCTELERIFGFP	SNMGRGARQKLLGRSWSVPVIRHLF
Dnmt3b	DVLWCTELERIFGFP	SNMGRGARQKLLGRSWSVPVIRHLF
Zmt3	DHIWITELEKIFGFP	KSMGRPORQRVLGKSWSVPVIRHLL
consensus	------E--R--GFP	----------R--G----P------

FIG. 5A

AMENDED SHEET

```
DNMT3A       EDICISCG......SLNVTLEHPLFVGGMCQNCKNCFLECAYQYDDDGYQSYCT
Dnmt3a       EDICISCG......SLNVTLEHPFFIGGMCQNCKNCFLECAYQYDDDGYQSYCT
DNMT3B       EDGCLSCG......RKNPVSFHPLFEGGLCQTCRDRFLELFYMYDDDGYQSYCT
Dnmt3b       EDRCLSCG......KKNPVSFHPLFEGGLCQSCRDRFLELFYMYDEDGYQSYCT
Zmt3         EDFCLSCG......SMSVDIIHPLFEGKLCTNCKFNFTETLYRYDEDGYQSYCT
ATRX_Human   IVSCTACGQQVNHFQKDSIYRHPSLQVLICKNCFKYYMSDDISRDSDGMDEQCR
ATRX_Mouse   IVSCTACGQQVNHFQKDSIYRHPSLQVLICKNCFKYYMSDDISRDSDGMDEQCR
Consensus       C  C                       C  C                   C

DNMT3A       ICCGGREVLMCGNNNCCRCFCVECVDLLVGPGAAQAAIKE.DPWNCYMCGHKGT
Dnmt3a       ICCGGREVLMCGNNNCCRCFCVECVDLLVGPGAAQAAIKE.DPWNCYMCGHKGT
DNMT3B       VCCEGRELLLCSNTSCCRCFCVECLEVLVGTGTAAEAKLQ.EPWSCYMCLPQRC
Dnmt3b       VCCEGRELLLCSNTSCCRCFCVECLEVLVGAGTAEDAKLQ.EPWSCYMCLPQRC
Zmt3         VCCSGMEVILCAHDSCCRSFCVDCLDILVCQGTFDRLKNV.DPWTCYLCAPETS
ATRX_Human   WCAEGGNLICC..DFCHNAFCKKCILRNLGRKELSTIMDENNQWYCYICHPEPL
ATRX_Mouse   WCAEGGNLICC..DFCHNAFCKKCILRNLGRKELSTIMDENNQWYCYICHPEPL
Consensus     C        C    C     C  C                     C  C
```

FIG.5B

AMENDED SHEET

Printed 20-06-2000

09-06-2000

PCT/US99/14373

09972006

AMENDED SHEET

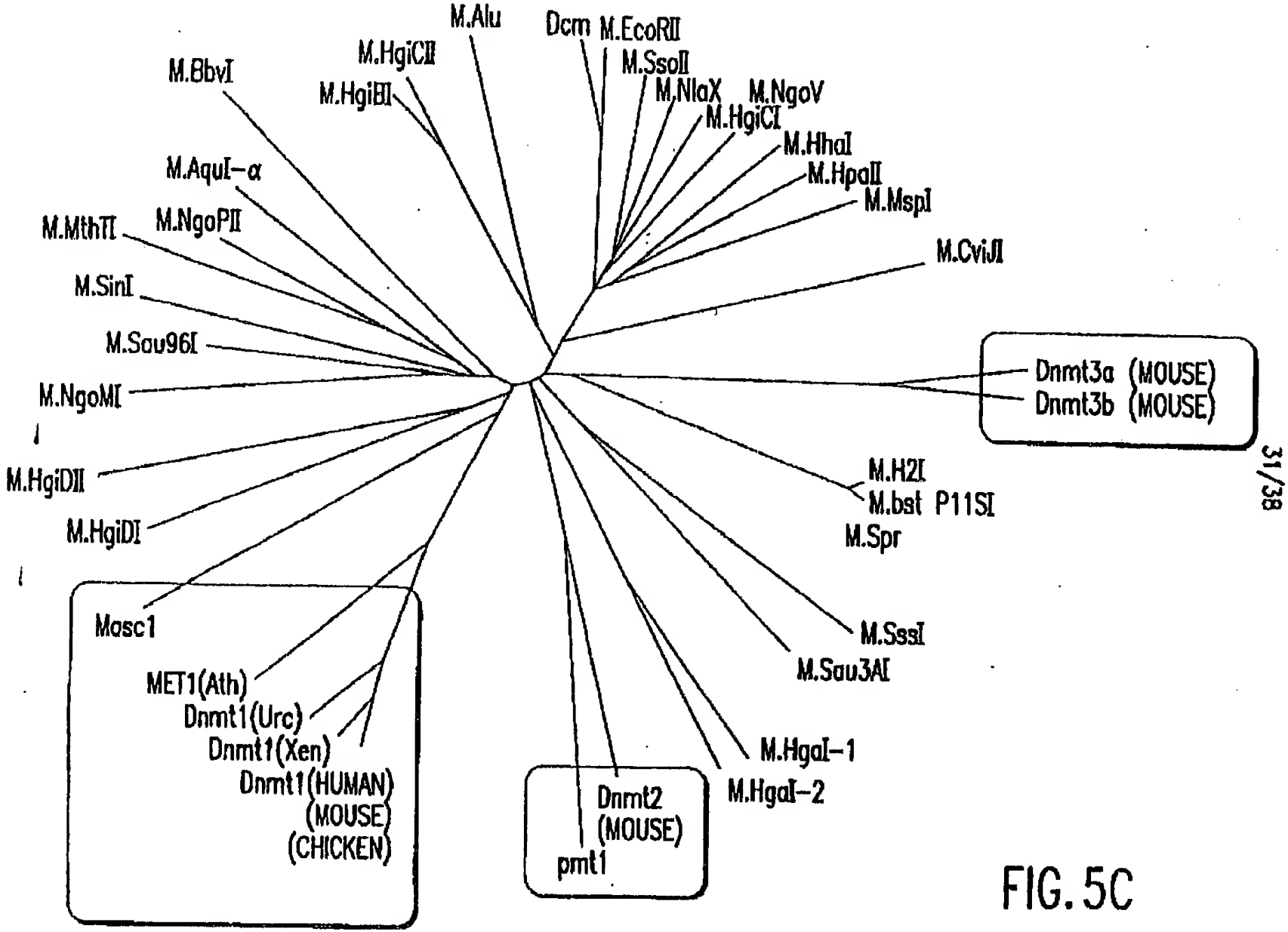

FIG. 5C

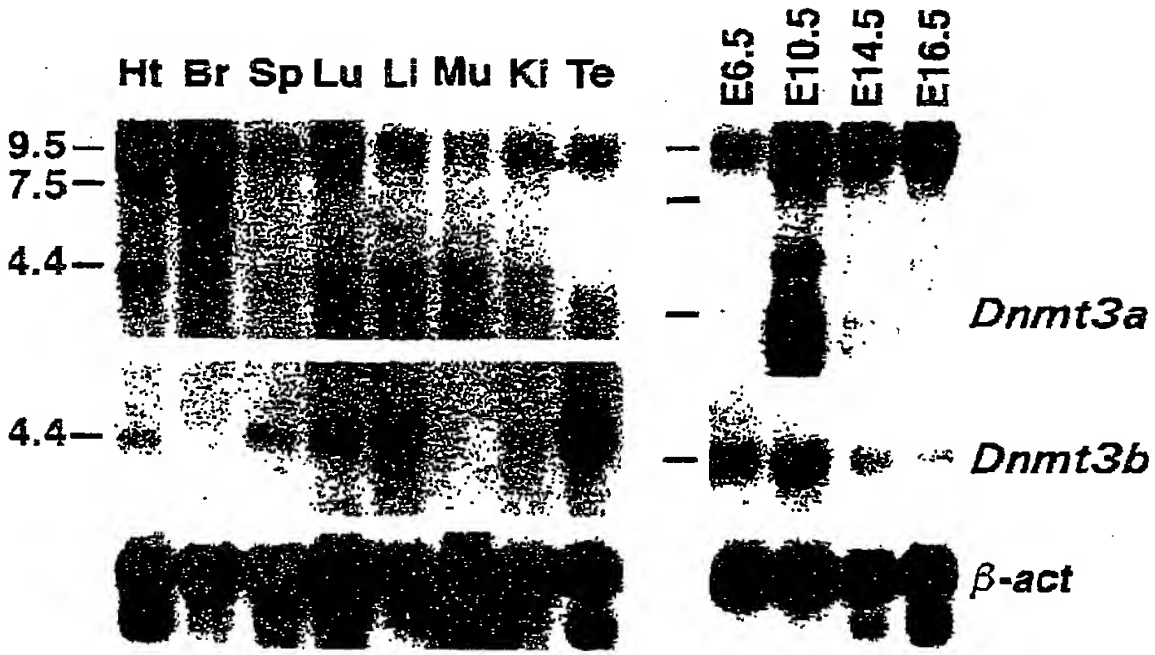

FIG.6A

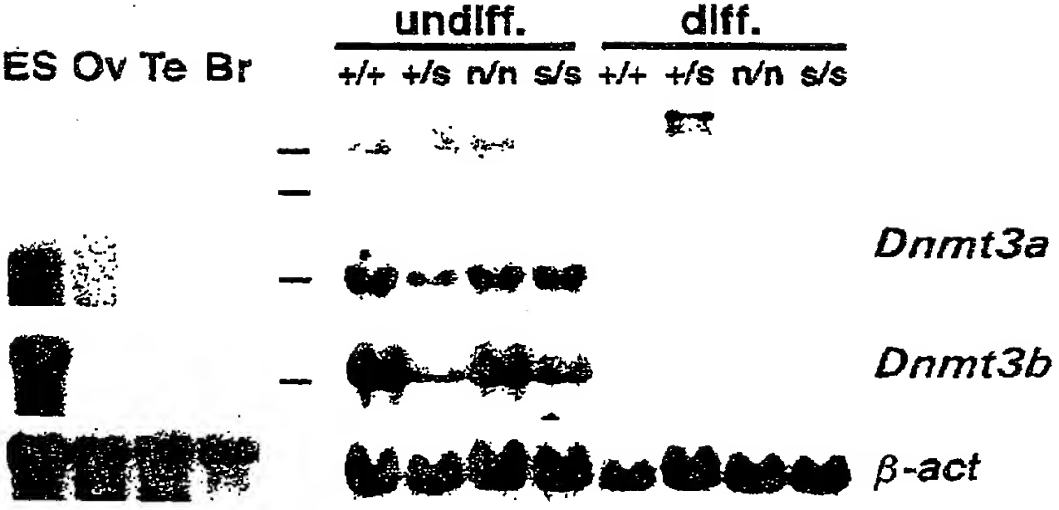

FIG.6B FIG.6C

AMENDED SHEET

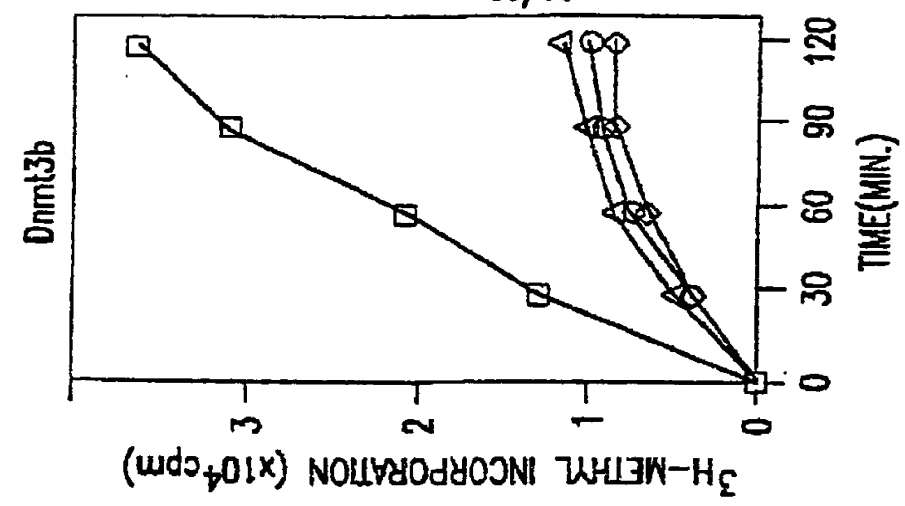

FIG. 7C

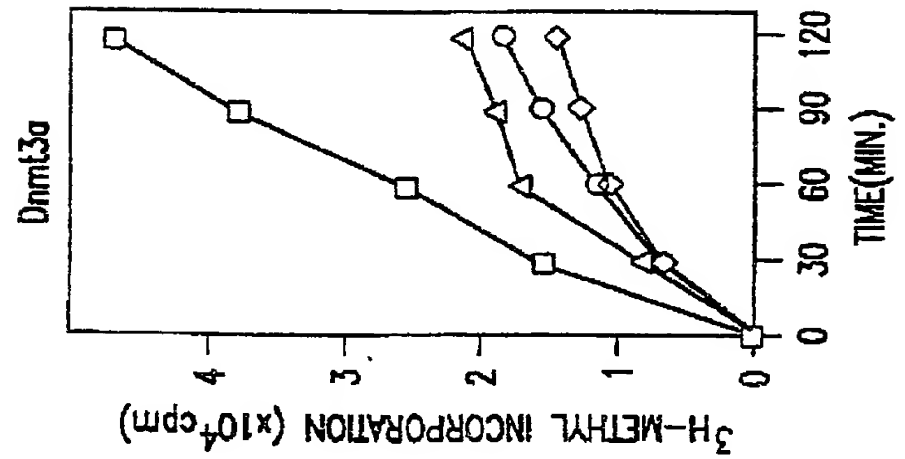

FIG. 7B

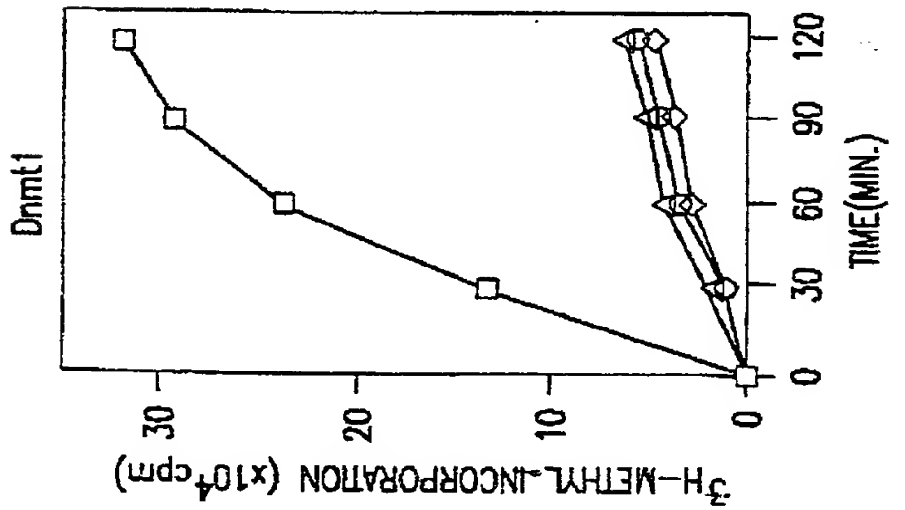

FIG. 7A

AMENDED SHEET

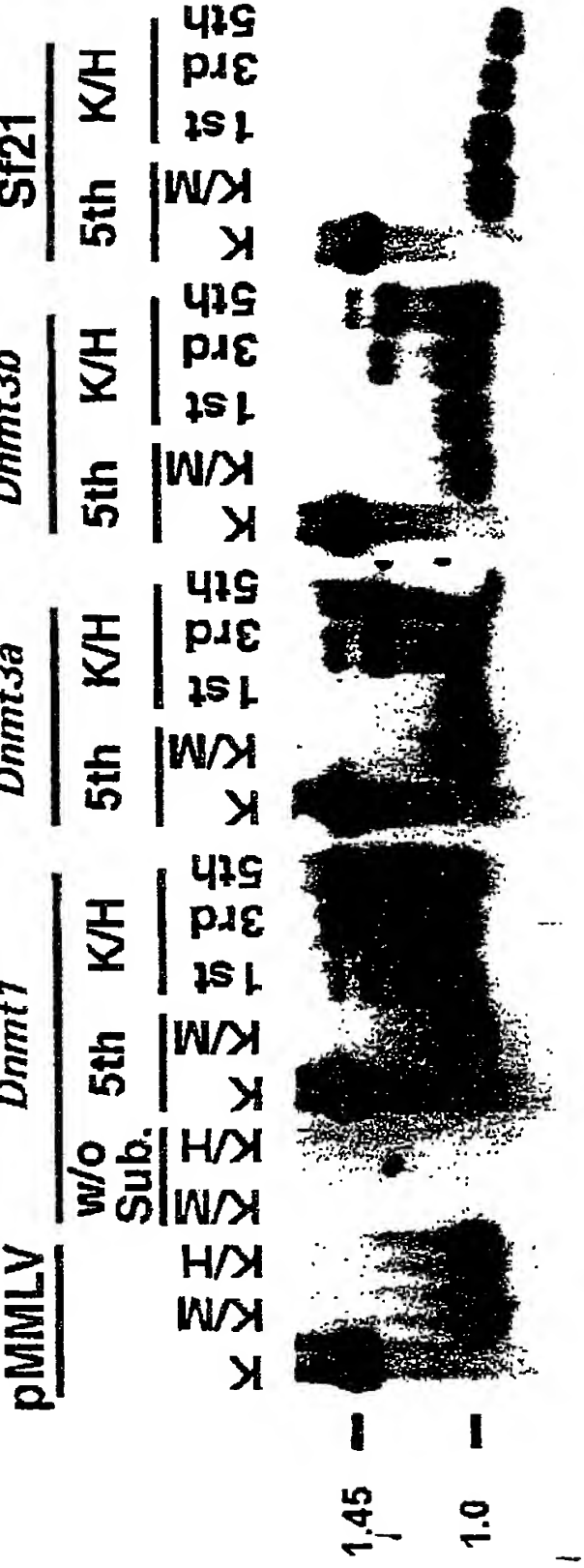
pMMLV
Dnmt1
Dnmt3a
Dnmt3b
Sf21
w/o Sub.
5th
1st
3rd
5th
K
K/M
K/H
1.45
1.0

FIG.7D

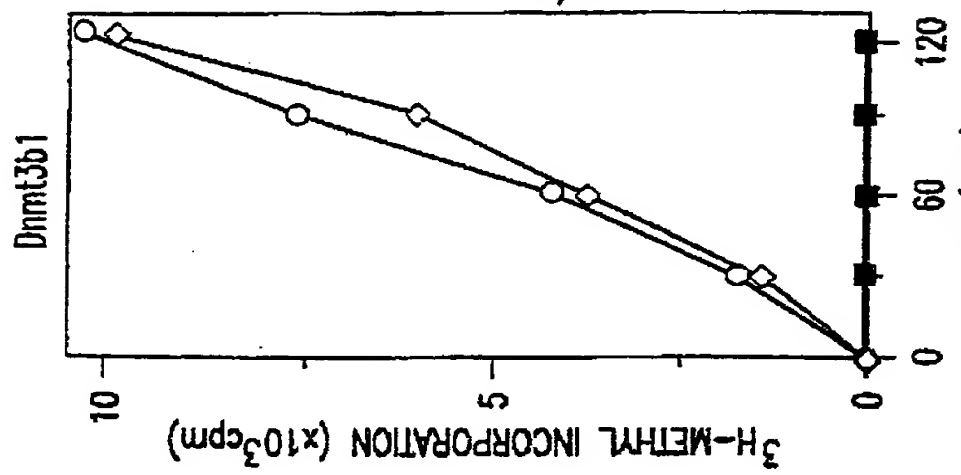

FIG. 8C

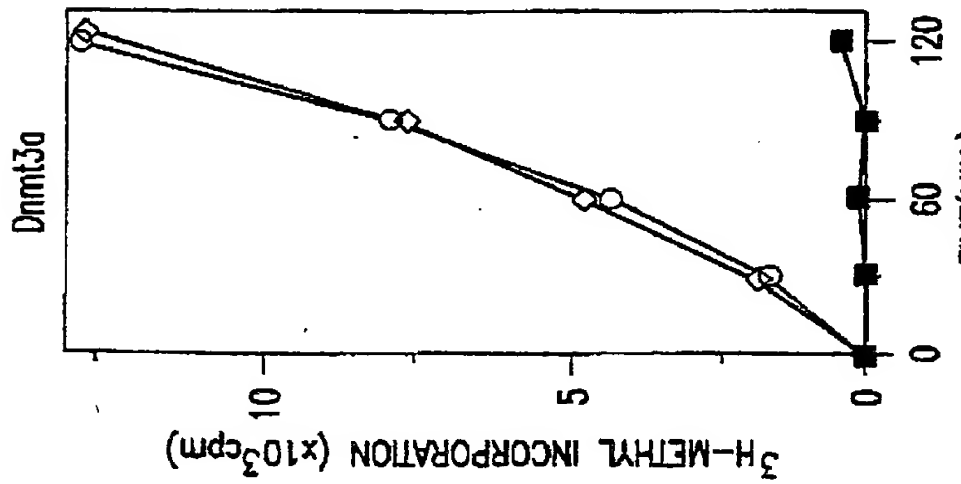

FIG. 8B

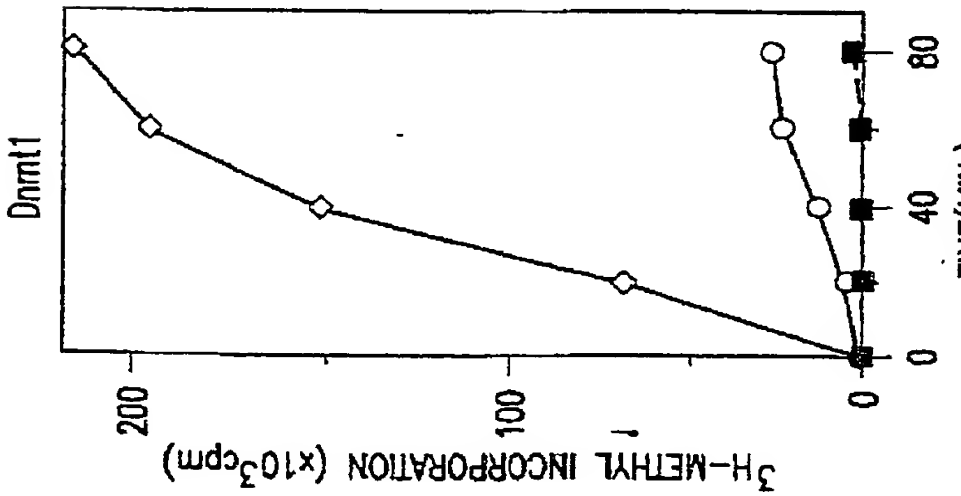

FIG. 8A

AMENDED SHEET

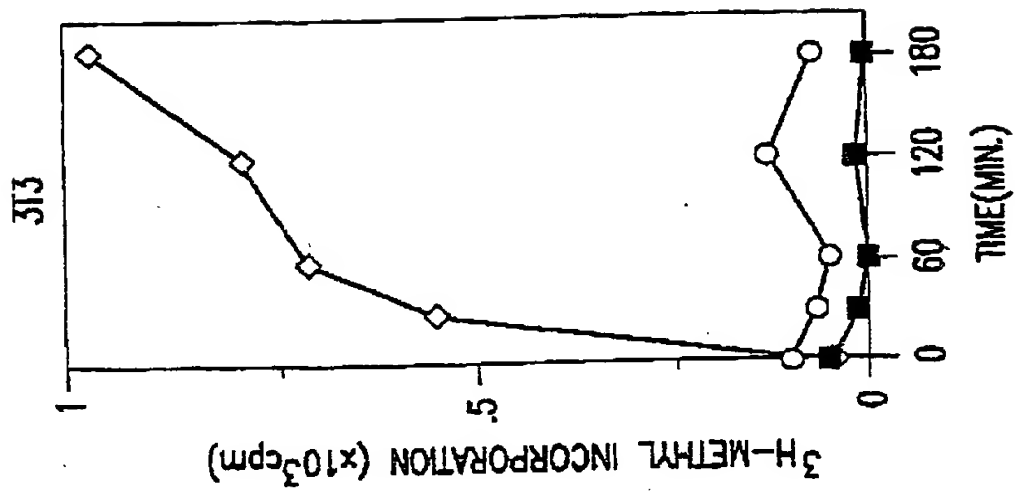

FIG. 8E

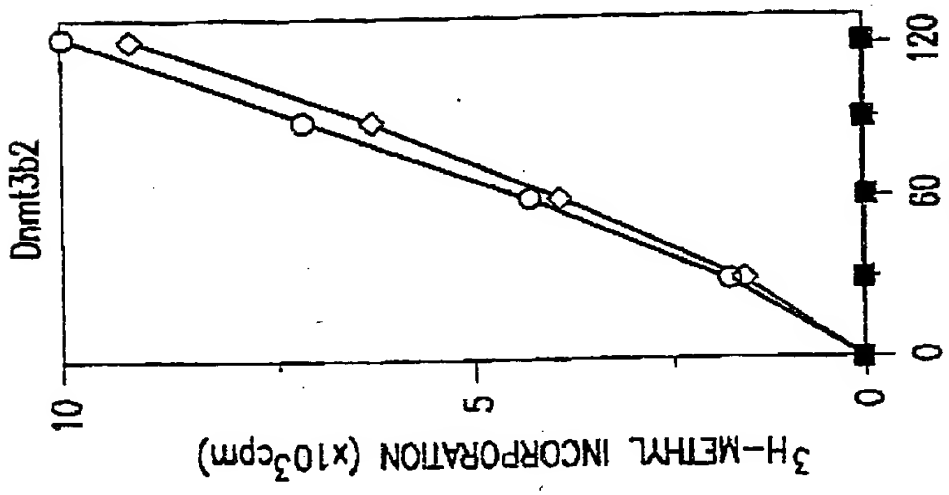

FIG. 8D

AMENDED SHEET

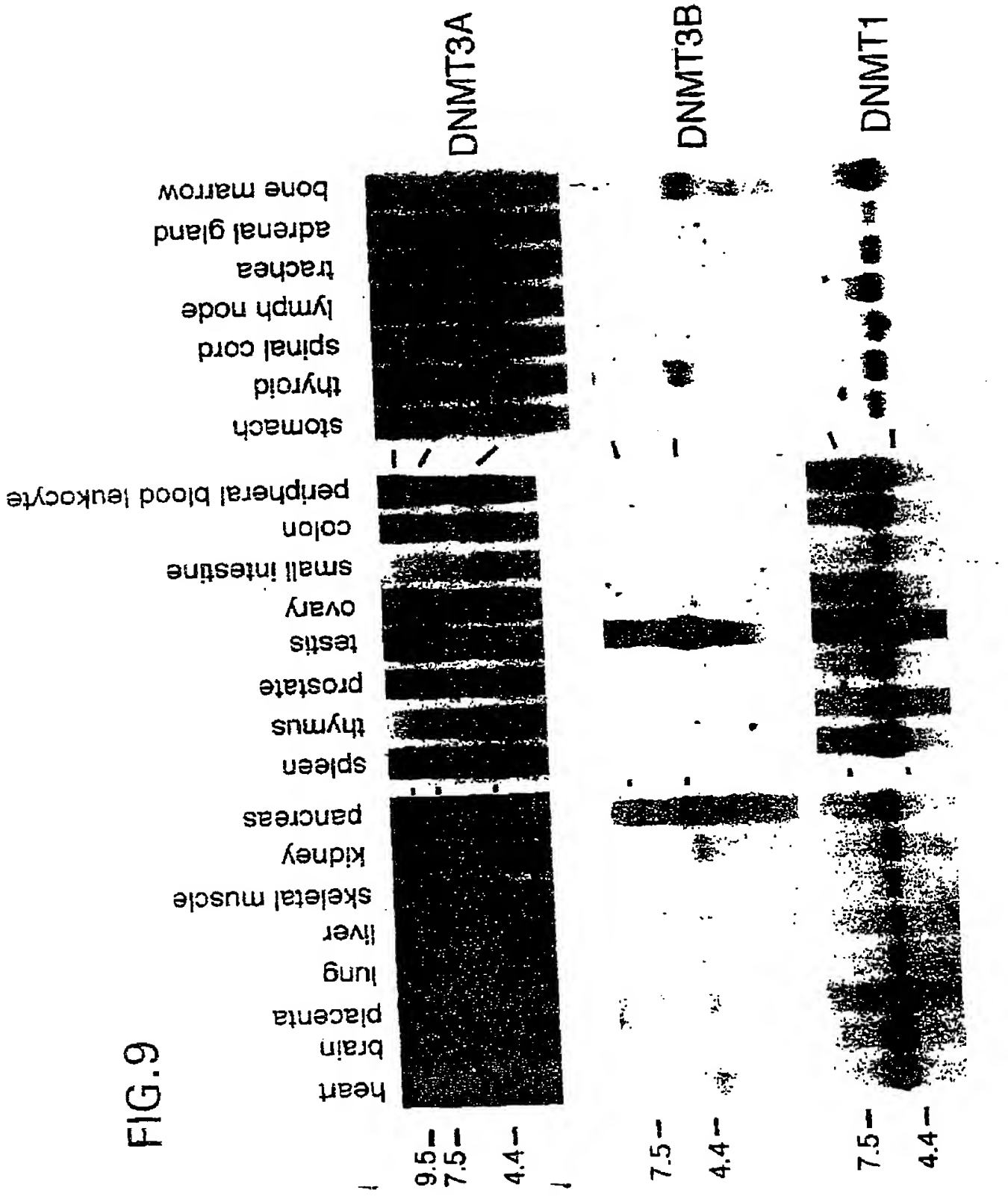

FIG.9

FOOTER "09806760"

promyelocytic leukamia HL-60
Hela cell S3
chronic myelogenous leukemia K-562
lymphoblastic leukemia MOLT-4
Burkitt's lymphoma Raji
colorectal adenocarcinoma SW480
lung carcinoma A549
melanoma G361

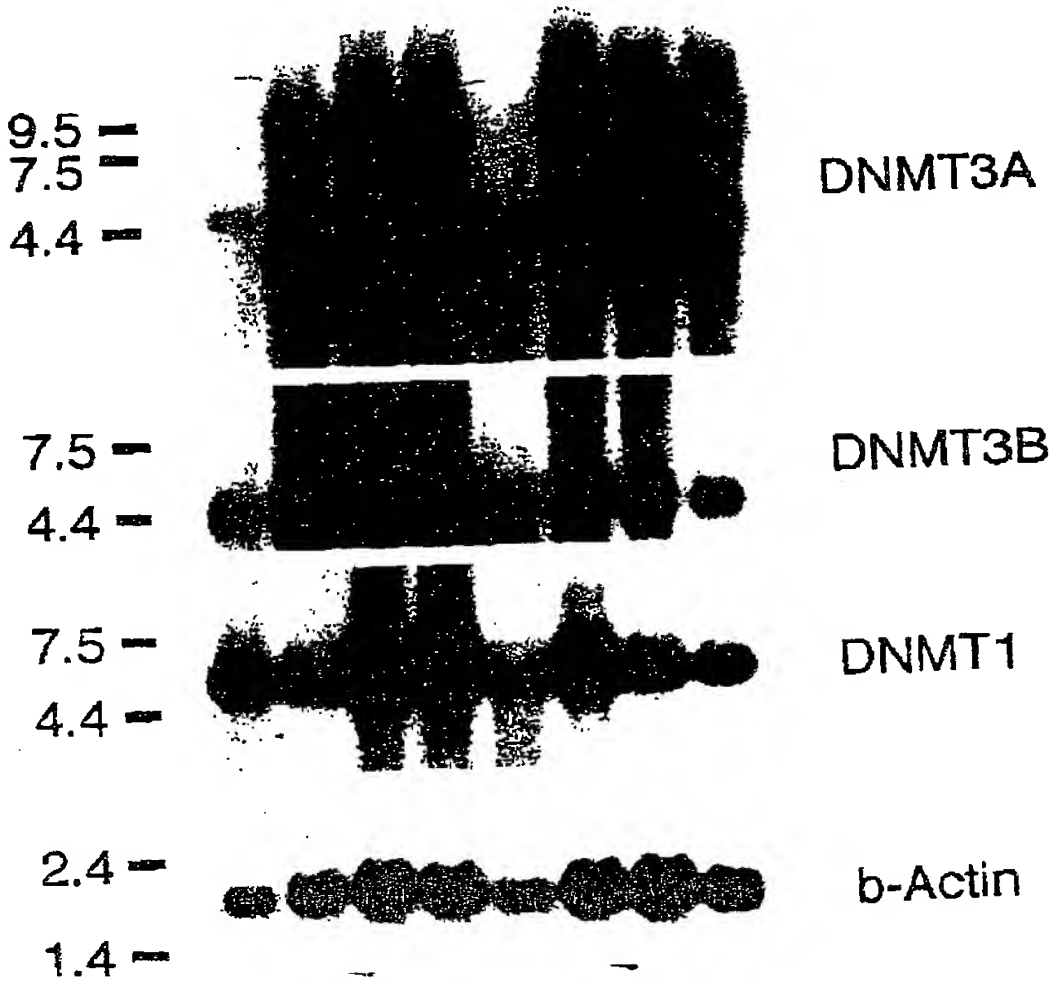

FIG.10

TOETZ0" 980024760